Relational Measures and Integration

Gunther Schmidt*

Institute for Software Technology, Department of Computing Science
Universität der Bundeswehr München, 85577 Neubiberg, Germany
Gunther.Schmidt@UniBw.DE

Abstract. Work in fuzzy modeling has recently made its way from the interval $[0, 1] \subseteq \mathbb{R}$ to the ordinal or even to the qualitative level. We proceed further and introduce relational measures and relational integration. First ideas of this kind, but for the real-valued linear orderings stem from Choquet (1950s) and Sugeno (1970s). We generalize to not necessarily linear order and handle it algebraically and in a componentfree manner. We thus open this area of research for treatment with theorem provers which would be extremely difficult for the classical presentation of Choquet and Sugeno integrals.

Keywords: Sugeno integral, Choquet integral, relation algebra, evidence and belief, plausibility, necessity, and possibility measures, relational measure.

1 Introduction

Mankind has developed a multitude of concepts to reason about something that is *better than* or is *more attractive than* something else or *similar to* something else. Such concepts lead to an enormous bulk of formulae and interdependencies.

We start from the concept of an *order* and a *strictorder*, defined as a transitive, antisymmetric, reflexive relation or as a transitive and asymmetric, respectively. In earlier times it was not at all clear that orderings need not be *linear* orderings. But since the development of lattice theory in the 1930s it became more and more evident that most of our reasoning with orderings was also possible when they failed to be linear ones. So the people studied fuzziness mainly along the linear order of $\mathbb{R}$ and began only later to generalize to the ordinal level: Numbers indicate the relative position of items, but no longer the magnitude of difference. Then they moved to the interval level: Numbers indicate the magnitude of difference between items, but there is no absolute zero point. Examples are attitude scales and opinion scales. We proceed even further and introduce relational measures with values in a lattice. Measures traditionally provide a basis for integration. Astonishingly, this holds true for these relational measures so that it becomes possible to introduce a concept of relational integration.

* Cooperation and communication around this research was partly sponsored by the European Cost Action 274: Tarski (*Theory and Applications of Relational Structures as Knowledge Instruments*), which is gratefully acknowledged.

R.A. Schmidt (Ed.): RelMiCS /AKA 2006, LNCS 4136, pp. 343–357, 2006.

2 Modelling Preferences

Who is about to make severe decisions will usually base these on carefully selected basic information and clean lines of reasoning. It is in general not too difficult to apply *just one* criterion and to operate according to this criterion. If several criteria must be taken into consideration, one has also to consider the all too often occurring situation that these provide contradictory information: "This car looks nicer, but it is much more expensive". Social and economical sciences have developed techniques to model what takes place when decisions are to be made in an environment with a multitude of diverging criteria. Preference is assumed to represent the degree to which one alternative is preferred to another. Often it takes the form of expressing that alternative A is considered being "not worse than" alternative B. Sometimes a linear ranking of the set of alternatives is assumed, which we avoid.

So finding decisions became abstracted to a scientific task. We may observe two lines of development. The Anglo-Saxon countries, in particular, formulated *utility theory*, in which numerical values shall indicate the intensity of some preference. Mainly in continental Europe, on the other hand side, binary relations were used to model pairwise preference; see [1], e.g. While the former idea allows to easily relate to statistics, the latter is based on evidence via direct comparison. In earlier years indeed, basic information was quite often statistical in nature and expressed in real numbers. Today we have more often fuzzy, vague, rough, etc. forms of qualification.

3 Introductory Example

We first give an example of relational integration deciding for a car to be bought out of several offers. We intend to follow a set $\mathcal{C}$ of three criteria, namely color, price, and speed. They are, of course, not of equal importance for us; price will most certainly outweigh the color of the car, e.g. Nevertheless let the valuation with these criteria be given on an ordinal scale $\mathcal{L}$ with 5 linearly ordered values as indicated on the left side of (1). (Here for simplicity, the ordering is linear, but it need not.) We name these values 1,2,3,4,5, but do not combine this with any arithmetic; i.e., value 4 is not intended to mean two times as good as value 2. Rather they might be described with linguistic variables as *bad, not totally bad, medium, outstanding, absolutely outstanding*; purposefully these example qualifications have not been chosen "equidistant".

$$\begin{array}{l} \text{color} \\ \text{price} \\ \text{speed} \end{array} \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \qquad \begin{aligned} 4 = \mathtt{lub}\,\big[\, & \mathtt{glb}\,(4_{v(color)}, 4_{\mu\{c,p\}}), \\ & \mathtt{glb}\,(4_{v(price)}, 4_{\mu\{c,p\}}), \\ & \mathtt{glb}\,(2_{v(speed)}, 5_{\mu\{c,p,s\}})\big] \end{aligned} \tag{1}$$

First we concentrate on the left side of (1). The task is to arrive at *one* overall valuation of the car out of these three. In a simple-minded approach, we might indeed conceive numbers $1, 2, 3, 4, 5 \in \mathbb{R}$ and then evaluate in a classical way the average value as $\frac{1}{3}(4 + 4 + 2) = 3.3333\ldots$, which is a value not expressible in the given scale. When considering the second example (2), we would arrive at the same average value although the switch from (1) to (2) between price and speed would trigger most people to decide differently.

$$\begin{array}{l} \text{color} \\ \text{price} \\ \text{speed} \end{array} \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \end{pmatrix} \qquad \begin{aligned} 3 = \texttt{lub}\,\big[\, & \texttt{glb}\,(4_{v(color)}, 3_{\mu\{c,s\}}), \\ & \texttt{glb}\,(2_{v(price)}, 5_{\mu\{c,p,s\}}), \\ & \texttt{glb}\,(4_{v(speed)}, 3_{\mu\{c,s\}})\big] \end{aligned} \tag{2}$$

With relational integration, we learn to make explicit which set of criteria to apply with which weight. It is conceivable that criteria c_1, c_2 are given a low weight but the criteria set $\{c_1, c_2\}$ in conjunction a high one. This means that we introduce a **relational measure** assigning values in $\mathcal{L}$ to subsets of $\mathcal{C}$.

$$\mu = \begin{array}{l} \{\} \\ \{\text{color}\} \\ \{\text{price}\} \\ \{\text{color,price}\} \\ \{\text{speed}\} \\ \{\text{color,speed}\} \\ \{\text{price,speed}\} \\ \{\text{color,price,speed}\} \end{array} \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix}$$

For gauging purposes we demand that the empty criteria set gets assigned the least value in $\mathcal{L}$ and the full criteria set the greatest. A point to stress is that we assume the criteria themselves as well as the measuring of subsets of criteria as *commensurable.*

The relational measure μ should obviously be monotonic with respect to the ordering Ω on the powerset of $\mathcal{C}$ and the ordering E on $\mathcal{L}$. We do not demand continuity (additivity), however. The price alone is ranked of medium importance 3, higher than speed alone, while color alone is considered completely unimportant and ranks 1. However, color and price together are ranked 4, i.e., higher than the supremum of ranks for color alone and for price alone, etc.

As now the valuations according to the criteria as well as the valuation according to the relative measuring of the criteria are given, we may proceed as visualized on the right sides of (1) and (2). We run through the criteria and always look for two items: their corresponding value and in addition for the value of that subset of criteria *assigning equal or higher values.* Then we determine the greatest lower bound for the two values. From the list thus obtained, the least upper bound is taken. The two examples above show how by simple evaluation along this concept, one will arrive at the overall values 4 or 3, respectively. This results from the fact that in the second case only such rather unimportant criteria as color and speed assign the higher values.

The effect is counterrunning: Low values of criteria as for s in (1) are intersected with rather high μ's as many criteria give higher scores and μ is monotonic. Highest values of criteria as for color or speed in (2) are intersected with the μ of a small or even one-element criteria set; i.e., with a rather small one. In total we find that here are two operations applied in a way we already know from matrix multiplication: a "sum" operator, `lub` or $\vee$, following application a "product" operator, `glb` or $\wedge$.

This example gave a first idea of how relational integration works and how it may be useful. Introducing a relational measure and using it for integration serves an important purpose: Concerns are now separated. One may design the criteria and the measure in a design phase prior to polling. Only then shall the questionnaire be filled, or the voters be polled. The procedure of coming to an overall valuation is now just computation and should no longer lead to quarrels.

4 Order-Theoretic Functionals

Given the page limit, we cannot present all the prerequisites on relation algebra and give [2,3] as a general reference for handling relations as boolean matrices and subsets of a set as boolean vectors. Let an order relation E be given on a set V. An element e is called an *upper bound* (also: *majorant*) of the subset of V characterized by the vector u of V provided $\forall x \in u : E_{xe}$. From the predicate logic version, we easily derive a relation-algebraic formulation as $e \subseteq \overline{\overline{E}^{\mathsf{T}} ; u}$, so that we introduce the order-theoretic functional $\mathtt{ubd}_E(u) := \overline{\overline{E}^{\mathsf{T}} ; u}$ to return the possibly empty vector of all upper bounds. Analogously, we have the set of lower bounds $\mathtt{lbd}_E(u) := \overline{\overline{E} ; u}$.

Starting herefrom, also the other traditional functionals may be obtained, as the least upper bound u, (also: *supremum*), the at most 1-element set of least elements among the set of all upper bounds of u

$$\mathtt{lub}_E(u) = \mathtt{ubd}_E(u) \cap \mathtt{lbd}_E(\mathtt{ubd}_E(u))$$

In contrast to our expectation that a least upper bound may exist or not, it will here always exist as a vector; it may, however be the null vector resembling that there is none.

As a tradition, a vector is often a column vector. In many cases, however, a row vector would be more convenient. We decided to introduce a variant denotation for order-theoretic functionals working on row vectors:

$$\mathtt{lubR}_E(X) := [\mathtt{lub}_E(X^{\mathsf{T}})]^{\mathsf{T}}, \text{ etc.}$$

We are here concerned with lattice orderings E only. For convenience we introduce notation for least and greatest elements as

$$0_E = \mathtt{glb}_E(\mathbb{T}), \quad 1_E = \mathtt{lub}_E(\mathbb{T})$$

5 Relational Measures

Assume the following basic setting with a set $\mathcal{C}$ of so-called criteria and a measuring lattice $\mathcal{L}$. Depending on the application envisaged, the set $\mathcal{C}$ may also be interpreted as one of players in a cooperative game, of attributes, of experts, or of voters in an opinion polling problem. This includes the setting with $\mathcal{L}$

the interval $[0, 1] \subseteq \mathbb{R}$ or a linear ordering for measuring. We consider a (relational) measure generalizing the concept of a fuzzy measure (or capacité in French origin) assigning measures in $\mathcal{L}$ for subsets of $\mathcal{C}$.

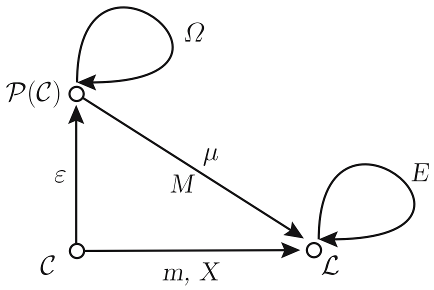

Fig. 5.1. Basic situation for relational integration

The relation ε is the membership relation between $\mathcal{C}$ and its powerset $\mathcal{P}(\mathcal{C})$. The measures envisaged will be called μ, other relations will be denoted as M. Valuations according to the criteria will be X or m depending on the context.

For a running example assume the task to assess persons of the staff according to their intellectual abilities as well as according to the workload they achieve to master.

$E =$

	(low,lazy)	(medium,lazy)	(low,fair)	(high,lazy)	(medium,fair)	(low,good)	(high,fair)	(medium,good)	(low,bulldozer)	(high,good)	(medium,bulldozer)	(high,bulldozer)
(low,lazy)	1	1	1	1	1	1	1	1	1	1	1	1
(medium,lazy)	0	1	0	1	1	0	1	1	0	1	1	1
(low,fair)	0	0	1	0	1	1	1	1	1	1	1	1
(high,lazy)	0	0	0	1	0	0	1	0	0	1	0	1
(medium,fair)	0	0	0	0	1	0	1	1	0	1	1	1
(low,good)	0	0	0	0	0	1	0	1	1	1	1	1
(high,fair)	0	0	0	0	0	0	1	0	0	1	0	1
(medium,good)	0	0	0	0	0	0	0	1	0	1	1	1
(low,bulldozer)	0	0	0	0	0	0	0	0	1	0	1	1
(high,good)	0	0	0	0	0	0	0	0	0	1	0	1
(medium,bulldozer)	0	0	0	0	0	0	0	0	0	0	1	1
(high,bulldozer)	0	0	0	0	0	0	0	0	0	0	0	1

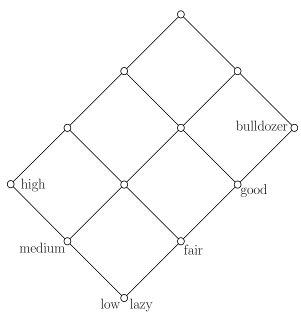

Fig. 5.2. Value lattice $\mathcal{L}$ ordered with E

5.1 Definition. Suppose a set of criteria $\mathcal{C}$ to be given together with some lattice $\mathcal{L}$, ordered by E, in which subsets of these criteria shall be given a measure

$\mu : \mathcal{P}(\mathcal{C}) \longrightarrow \mathcal{L}$. Let Ω be the ordering on $\mathcal{P}(\mathcal{C})$. We call a mapping $\mu : \mathcal{P}(\mathcal{C}) \rightarrow \mathcal{L}$ a **(relational) measure** provided

- $\Omega;\mu \subseteq \mu;E$, meaning that μ is isotonic wrt. to the orderings Ω and E.
- $\mu^\mathsf{T};0_\Omega = 0_E$, meaning that the empty subset of $\mathcal{P}(\mathcal{C})$ is mapped to the least element of $\mathcal{L}$.
- $\mu^\mathsf{T};1_\Omega = 1_E$, meaning that the full subset of $\mathcal{P}(\mathcal{C})$ is mapped to the greatest element of $\mathcal{L}$. □

A (relational) measure for $s \in \mathcal{P}(\mathcal{C})$, i.e., $\mu(s)$ when written as a mapping or $\mu^\mathsf{T};s$ when written in relation form, may be interpreted as the weight of importance we attribute to the combination s of criteria. It should not be mixed up with a probability. The latter would require the setting $\mathcal{L} = [0,1] \subseteq \mathbb{R}$ and in addition that μ be continuous.

Many ideas of this type have been collected by Glenn Shafer under the heading *theory of evidence*, calling μ a *belief function*. Using it, he explained a basis of rational behaviour. We attribute certain weights to evidence, but do not explain in which way. These weights shall in our case be lattice-ordered. This alone gives us reason to rationally decide this or that way. Real-valued belief functions have numerous applications in artificial intelligence, expert systems, approximate reasoning, knowledge extraction from data, and Bayesian Networks.

Concerning additivity, the example of Glenn Shafer [4] is when one is wondering whether a Ming vase is a genuine one or a fake. We have to put the full amount of our belief on the disjunction "*genuine* or *fake*" as one of the alternatives will certainly be the case. But the amount of trust we are willing to put on the alternatives may in both cases be very small as we have only tiny hints for being genuine, but also very tiny hints for being a fake.

With the idea of probability, we could not so easily cope with the ignorance just mentioned. Probability does not allow one to withhold belief from a proposition without *according the withheld amount of belief to the negation*. When thinking on the Ming vase in terms of probability we would have to attribute p to *genuine* and $1-p$ to *fake*.

In the extreme case, we have complete ignorance expressed by the so-called **vacuous belief mapping**

$$\mu_0(s) = \begin{cases} 0_E & \text{if } \mathcal{C} \neq s \\ 1_E & \text{if } \mathcal{C} = s \end{cases}$$

On the other side, we may completely overspoil our trust expressed by what we may call a **light-minded belief mapping**

$$\mu_1(s) = \begin{cases} 0_E & \text{if } 0_\Omega = s \\ 1_E & \text{otherwise} \end{cases}$$

To an arbitrary non-empty set of criteria, the light-minded belief mapping attributes all the components of trust or belief.

5.2 Definition. Given this setting, we call μ

i) a **Bayesian measure** if it is lattice-continuous, i.e.,

$$\texttt{lub}_E(\mu^\top; s) = \mu^\top; \texttt{lub}_\Omega(s)$$

for a subset $s \subseteq \mathcal{P}(\mathcal{C})$, or expressed differently, a set of subsets of $\mathcal{C}$.

ii) a **simple support mapping** focused on U valued with v, if U is a non-empty subset $U \subseteq \mathcal{C}$ and $v \in \mathcal{L}$ an element such that

$$\mu(s) = \begin{cases} 0_E & \text{if } s \not\supseteq U \\ v & \text{if } \mathcal{C} \neq s \supseteq U \\ 1_E & \text{if } \mathcal{C} = s \end{cases}$$ □

In particular, μ_1 is Bayesian while μ_0 is not. In the real-valued environment, the condition for a Bayesian measure is: additive when non-overlapping. Lattice-continuity incorporates two concepts, namely additivity

$$\mu^\top; (s_1 \cup s_2) = \mu^\top; s_1 \ \cup_\mathcal{L} \ \mu^\top; s_2$$

and sending 0_Ω to 0_E.

Combining measures

Dempster [5] found for the real-valued case a way of combining measures in a form closely related to conditional probability. It shows a way of adjusting opinion in the light of new evidence. We have re-modeled this for the relational case. One should be aware of how a measure behaves on upper and lower cones:

$$\mu = \texttt{lubR}_E(\Omega^\top; \mu) \quad \mu = \texttt{glbR}_E(\Omega; \mu)$$

When one has in addition to μ got further evidence from a second measure μ', one will intersect the upper cones resulting in a possibly smaller cone positioned higher up and take its greatest lower bound:

$$\mu \oplus \mu' := \texttt{glbR}_E(\mu; E \cap \mu'; E)$$

One might, however, also look where μ and μ' agree, and thus intersect the lower bound cones resulting in a possibly smaller cone positioned deeper down and take its least upper bound:

$$\mu \otimes \mu' := \texttt{lubR}_E(\mu; E^\top \cap \mu'; E^\top)$$

5.3 Proposition. If the measures μ, μ' are given, $\mu \oplus \mu'$ as well as $\mu \otimes \mu'$ are measures again. Both operations are commutative and associative. The vacuous belief mapping μ_0 is the null element while the light-minded belief mapping μ_1 is the unit element among measures:

$$\mu \oplus \mu_0 = \mu, \quad \mu \otimes \mu_1 = \mu, \quad \text{and} \quad \mu \otimes \mu_0 = \mu_0$$

Proof: The least element must be sent to the least element. This result is prepared observing that 0_Ω is a transposed mapping, in

$\texttt{lbd}_E([\mu; E \cap \mu'; E]^\top); 0_\Omega$
$= \overline{\overline{E}; [\mu; E \cap \mu'; E]^\top}; 0_\Omega$
$= \overline{\overline{E}; [\mu; E \cap \mu'; E]^\top; 0_\Omega}$ a mapping may slip under a negation from the left

$= \overline{\overline{E}; [E^\mathsf{T}; \mu^\mathsf{T} \cap E^\mathsf{T}; \mu'^\mathsf{T}]; 0_\Omega}$
$= \overline{\overline{E}; [E^\mathsf{T}; \mu^\mathsf{T}; 0_\Omega \cap E^\mathsf{T}; \mu'^\mathsf{T}; 0_\Omega]}$ multiplying an injective relation from the right
$= \overline{\overline{E}; [E^\mathsf{T}; 0_E \cap E^\mathsf{T}; 0_E]}$ definition of measure
$= \overline{\overline{E}; E^\mathsf{T}; 0_E}$
$= \overline{\overline{E}; \mathbb{T}}$ in the complete lattice E
$= \mathtt{lbd}(\mathbb{T}) = 0_E$ in the complete lattice E

Now
$(\mu \oplus \mu')^\mathsf{T}; 0_\Omega = \mathtt{glb}_E([\mu; E \cap \mu'; E]^\mathsf{T}); 0_\Omega$
$= \big(\mathtt{lbd}_E([\mu; E \cap \mu'; E]^\mathsf{T}) \cap \mathtt{ubd}(\mathtt{lbd}_E([\mu; E \cap \mu'; E]^\mathsf{T}))\big); 0_\Omega$
$= \mathtt{lbd}_E([\mu; E \cap \mu'; E]^\mathsf{T}); 0_\Omega \cap \overline{\overline{E}^\mathsf{T}; \mathtt{lbd}_E([\mu; E \cap \mu'; E]^\mathsf{T}); 0_\Omega}$
$= 0_E \cap \overline{\overline{E}^\mathsf{T}; \mathtt{lbd}_E([\mu; E \cap \mu'; E]^\mathsf{T}); 0_\Omega}$
$= 0_E \cap \overline{\overline{E}^\mathsf{T}; 0_E}$
$= 0_E \cap \mathtt{ubd}(0_E)$
$= 0_E \cap \mathbb{T} = 0_E$

For reasons of space, the other parts of the proof are left to the reader. □

6 Relational Integration

Assume now that for all the criteria $\mathcal{C}$ a valuation has taken place resulting in a mapping $X : \mathcal{C} \longrightarrow \mathcal{L}$. The question is how to arrive at an overall valuation by rational means, for which μ shall be the guideline.

6.1 Definition. Given a relational measure μ and a mapping X indicating the values given by the criteria, we define the **relational integral**

$$(R) \int X \circ \mu \; := \mathtt{lubR}_E(\mathbb{T}; \mathtt{glbR}_E[(X \cup \mathtt{syq}(X; E; X^\mathsf{T}, \varepsilon); \mu)]) \qquad \square$$

As already mentioned, we apply a sum operator `lub` after applying the product operator `glb`. When values are assigned with X, we look with E for those greater or equal, then with X^T for the criteria so valuated. Now comes a technically difficult step, namely proceeding to the union of the resulting sets with the symmetric quotient `syq` and the membership relation ε. The μ-score of this set is then taken.

The tables in Fig. 6.1 show a measure, a valuation and then the relational integral computed with the TituRel system.

We are now in a position to understand why gauging $\mu^\mathsf{T}; 1_\Omega = 1_E$ is necessary for μ, or "greatest element is sent to greatest element". Consider, e.g., the special case of an X with all criteria assigning the same value. We certainly expect the relational integral to precisely deliver this value regardless of the measure chosen. But this might not be the case if a measure should assign too small a value to the full set.

$\mu =$

	(low,lazy)	(medium,lazy)	(low,fair)	(high,lazy)	(medium,fair)	(low,good)	(high,fair)	(medium,good)	(low,bulldozer)	(high,good)	(medium,bulldozer)	(high,bulldozer)
{}	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe}	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Bob}	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe,Bob}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Carl}	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe,Carl}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Bob,Carl}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**
{Abe,Bob,Carl}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Don}	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
{Bob,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
{Abe,Bob,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Carl,Don}	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe,Carl,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Bob,Carl,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Abe,Bob,Carl,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**

$X =$

Abe	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Bob	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
Carl	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Don	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

$$(R)\int X \circ \mu = (\mathbf{0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0})$$

Fig. 6.1. Measure, a valuation and the relational integral

These considerations originate from a free re-interpretation of the following concepts for work in $[0,1] \subseteq \mathbb{R}$. The **Sugeno integral** operator is in the literature defined as

$$M_{S,\mu}(x_1 \ldots, x_m) = (S)\int x \circ \mu = \bigvee_{i=1}^{m} [x_i \wedge \mu(A_i)]$$

and the **Choquet integral** operator as

$$M_{C,\mu}(x_1, \ldots, x_m) = (C)\int x \circ \mu = \sum_{i=1}^{m} [(x_i - x_{i-1}) \cdot \mu(A_i)]$$

In both cases the elements of vector $(x_1, \ldots, x_m)$, and parallel to this, the criteria set $\mathcal{C} = \{C_1, \ldots, C_m\}$ have each time been reordered such that

$$0 = x_0 \le x_1 \le x_2 \le \cdots \le x_m \le x_{m+1} = 1 \text{ and } \mu(A_i) = \mu(C_i, \ldots, C_m).$$

The concept of Choquet integral was first introduced for a real-valued context in [6] and later used by Michio Sugeno [7]. This integral has nice properties for aggregation: It is continuous, non-decreasing, and stable under certain interval preserving transformations. Not least reduces it to the weighted arithmetic mean as soon as it becomes additive.

7 Defining Relational Measures

Such measures may be given directly, which is, however, a costly task as a powerset is involved all of whose elements need values. Therefore, they mainly originate in some other way.

Measures originating from direct valuation of criteria

Let a **direct valuation** of the criteria be given as any relation m between $\mathcal{C}$ and $\mathcal{L}$. Although it is allowed to be contradictory and non-univalent, we provide for a way of defining a relational measure based on it. This will happen via the following constructs

$$\sigma(m) := \overline{\varepsilon^{\mathsf{T}} ; m ; \overline{E}} \qquad \pi(\mu) := \overline{\varepsilon ; \mu ; \overline{E}^{\mathsf{T}}}, \tag{3}$$

which very obviously satisfy the Galois correspondence requirement

$$m \subseteq \pi(\mu) \iff \mu \subseteq \sigma(m).$$

They satisfy $\sigma(m ; E^{\mathsf{T}}) = \sigma(m)$ and $\pi(\mu ; E) = \pi(\mu)$, so that in principle only lower, respectively upper, cones occur as arguments. Applying $\overline{W ; E} = \overline{W ; E} ; E^{\mathsf{T}}$, we get

$$\sigma(m) ; E = \overline{\varepsilon^{\mathsf{T}} ; m ; \overline{E}} ; E = \overline{\varepsilon^{\mathsf{T}} ; m ; \overline{E} ; E^{\mathsf{T}}} ; E = \overline{\varepsilon^{\mathsf{T}} ; m ; \overline{E}} = \sigma(m),$$

so that images of σ are always upper cones — and thus best described by their greatest lower bound $\mathtt{glbR}_E(\sigma(m))$.

7.1 Proposition. Given any relation $m : \mathcal{C} \to \mathcal{L}$, the construct

$$\mu_m := \mu_0 \oplus \mathtt{glbR}_E(\sigma(m))$$

forms a relational measure, the so-called **possibility measure**. □

Addition of the vacuous belief mapping μ_0 is again necessary for gauging purposes. In case m is a mapping, the situation becomes even nicer. From

$$\pi(\sigma(m ; E^{\mathsf{T}})) = \pi(\sigma(m)) = \overline{\varepsilon ; \overline{\varepsilon^{\mathsf{T}} ; m ; \overline{E}} ; \overline{E}^{\mathsf{T}}}$$
$$= \overline{\overline{m ; \overline{E}} ; \overline{E}^{\mathsf{T}}} \text{ as it can be shown that in general } \varepsilon ; \overline{\varepsilon^{\mathsf{T}} ; X} = \overline{X} \text{ for all } X$$
$$= m ; \overline{\overline{\overline{E}} ; \overline{E}^{\mathsf{T}}} \text{ as } m \text{ was assumed to be a mapping}$$
$$= m ; \overline{E ; \overline{E}^{\mathsf{T}}}$$
$$= m ; E^{\mathsf{T}}$$

we see that this is an adjunction on cones. The lower cones $m ; E^{\mathsf{T}}$ in turn are $1 : 1$ represented by their least upper bounds $\mathtt{lubR}_E(m ; E)$.

The following proposition exhibits that a Bayesian measure is a rather special case, namely more or less directly determined as a possibility measure for a direct valuation via a mapping m. Fig. 7.1 shows an example. One may proceed from m to the measure according to Prop. 7.1 or vice versa according to Prop. 7.2.

7.2 Proposition. Let μ be a Bayesian measure. Then $m_\mu := \mathtt{lubR}_E(\pi(\mu))$ is that direct valuation for which $\mu = \mu_{m_\mu}$. □

$\mu_B =$

	(low,lazy)	(medium,lazy)	(low,fair)	(high,lazy)	(medium,fair)	(low,good)	(high,fair)	(medium,good)	(low,bulldozer)	(high,good)	(medium,bulldozer)	(high,bulldozer)
{}	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe}	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Bob}	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
{Abe,Bob}	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
{Carl}	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe,Carl}	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
{Bob,Carl}	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
{Abe,Bob,Carl}	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
{Don}	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**
{Bob,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Abe,Bob,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Carl,Don}	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**
{Abe,Carl,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**
{Bob,Carl,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**
{Abe,Bob,Carl,Don}	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**

$m_{\mu_B} =$

Abe	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Bob	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**
Carl	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Don	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**

Fig. 7.1. Direct valuation with corresponding Bayesian measure

With this method just a few of the many relational measures will be found. By construction they are all continuous (or additive).

Measures originating from a body of evidence

We may also derive relational measures out of some relation between $\mathcal{P}(\mathcal{C})$ and $\mathcal{L}$. Although it is allowed to be non-univalent, we provide for a way of defining two measures based on it — which may coincide.

7.3 Definition. Let our general setting be given.

i) A **body of evidence** is an arbitrary relation $M : \mathcal{P}(\mathcal{C}) \longrightarrow \mathcal{L}$, restricted by the requirement that $M^\mathsf{T}; 0_\Omega \subseteq 0_E$.

ii) When the body of evidence M is in addition a mapping, we speak — following [4] — of a **basic probability assignment**. □

If I dare saying that occurrence of $A \subseteq \mathcal{C}$ deserves my trust to the amount $M(A)$, then $A' \subseteq A \subseteq \mathcal{C}$ deserves at least this amount of trusting as it occurs whenever A occurs. I might, however, not be willing to consider that $A'' \subseteq \mathcal{C}$ with $A \subseteq A''$ deserves to be trusted with the same amount as there is a chance that it occurs not so often.

$M :=$

	(low,lazy)	(medium,lazy)	(low,fair)	(high,lazy)	(medium,fair)	(low,good)	(high,fair)	(medium,good)	(low,bulldozer)	(high,good)	(medium,bulldozer)	(high,bulldozer)
{}	0	0	0	0	0	0	0	0	0	0	0	0
{Abe}	0	0	0	0	0	0	0	0	0	0	0	0
{Bob}	0	0	1	0	0	0	0	0	0	0	0	0
{Abe,Bob}	0	0	0	0	0	0	0	0	0	0	0	0
{Carl}	0	0	0	0	0	0	0	0	0	0	0	0
{Abe,Carl}	0	0	0	0	0	0	0	0	0	0	0	0
{Bob,Carl}	0	0	0	0	0	0	0	0	0	0	0	0
{Abe,Bob,Carl}	0	0	0	0	0	0	0	0	0	0	0	0
{Don}	0	0	0	0	0	0	0	0	0	0	0	0
{Abe,Don}	0	0	0	0	0	0	0	0	0	0	0	0
{Bob,Don}	0	0	0	0	0	0	0	0	0	0	0	0
{Abe,Bob,Don}	0	0	0	0	1	0	0	0	0	0	0	0
{Carl,Don}	0	0	0	0	0	0	0	0	0	0	0	0
{Abe,Carl,Don}	0	0	0	0	0	0	0	0	1	0	0	0
{Bob,Carl,Don}	0	0	0	0	0	0	0	0	0	0	0	0
{Abe,Bob,Carl,Don}	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 7.2. A body of evidence

We should be aware that the basic probability assignment is meant to assign something to a set regardless of what is assigned to its proper subsets. The condition $M^{\mathsf{T}}; 0_\Omega \subseteq 0_E$ expresses that M either does not assign any belief to the empty set or assigns it just 0_E.

Now a construction similar to that in (3) becomes possible, introducing

$$\sigma'(M) := \overline{\Omega^{\mathsf{T}}; M; \overline{E}} \qquad \pi'(\mu) := \overline{\Omega; \mu; \overline{E}^{\mathsf{T}}}, \tag{4}$$

which again satisfies the Galois correspondence requirement

$$M \subseteq \pi'(\mu) \iff \mu \subseteq \sigma'(M).$$

Obviously $\sigma'(M; E^{\mathsf{T}}) = \sigma'(M)$ and $\pi'(\mu; E) = \pi'(\mu)$, so that in principle only upper (E) and lower (E^{T}), respectively, cones are set into relation. But again applying $\overline{W; E} = \overline{W; E}; E^{\mathsf{T}}$, we get

$$\sigma'(M) ; E = \overline{\Omega^\mathsf{T} ; M ; \overline{E}} ; E = \overline{\Omega^\mathsf{T} ; M ; \overline{E} ; E^\mathsf{T}} ; E = \overline{\Omega^\mathsf{T} ; M ; \overline{E}} = \sigma'(M),$$

so that images of σ' are always upper cones — and thus best described by their greatest lower bound $\mathtt{glbR}_E(\sigma'(M))$.

7.4 Proposition. Should some body of evidence M be given, there exist two relational measures closely resembling M,

i) the **belief measure** $\mu_{\text{belief}}(M) := \mu_0 \oplus \mathtt{lubR}_E(\Omega^\mathsf{T} ; M)$
ii) the **plausibility measure** $\mu_{\text{plausibility}}(M) := \mu_0 \oplus \mathtt{lubR}_E((\Omega \cap \overline{\overline{\Omega} ; \mathbb{T}})^\mathsf{T} ; \Omega ; M)$
iii) In general, the belief measure assigns values not exceeding those of the plausibility measure, i.e., $\mu_{\text{belief}}(M) \subseteq \mu_{\text{plausibility}}(M) ; E^\mathsf{T}$. □

$\mu_{\text{belief}}(M)$

	(low,lazy)	(medium,lazy)	(low,fair)	(high,lazy)	(medium,fair)	(low,good)	(high,fair)	(medium,good)	(low,bulldozer)	(high,good)	(medium,bulldozer)	(high,bulldozer)
{}	1	0	0	0	0	0	0	0	0	0	0	0
{Abe}	1	0	0	0	0	0	0	0	0	0	0	0
{Bob}	0	0	1	0	0	0	0	0	0	0	0	0
{Abe,Bob}	0	0	1	0	0	0	0	0	0	0	0	0
{Carl}	1	0	0	0	0	0	0	0	0	0	0	0
{Abe,Carl}	1	0	0	0	0	0	0	0	0	0	0	0
{Bob,Carl}	0	0	1	0	0	0	0	0	0	0	0	0
{A,B,C}	0	0	1	0	0	0	0	0	0	0	0	0
{Don}	1	0	0	0	0	0	0	0	0	0	0	0
{Abe,Don}	1	0	0	0	0	0	0	0	0	0	0	0
{Bob,Don}	0	0	1	0	0	0	0	0	0	0	0	0
{A,B,D}	0	0	0	0	1	0	0	0	0	0	0	0
{Carl,Don}	1	0	0	0	0	0	0	0	0	0	0	0
{A,C,D}	0	0	0	0	0	0	0	0	1	0	0	0
{B,C,D}	0	0	1	0	0	0	0	0	0	0	0	0
{A,B,C,D}	0	0	0	0	0	0	0	0	0	0	0	1

$\mu_{\text{plausibility}}(M)$

	(low,lazy)	(medium,lazy)	(low,fair)	(high,lazy)	(medium,fair)	(low,good)	(high,fair)	(medium,good)	(low,bulldozer)	(high,good)	(medium,bulldozer)	(high,bulldozer)
{}	1	0	0	0	0	0	0	0	0	0	0	0
{Abe}	0	0	0	0	0	0	0	0	0	0	1	0
{Bob}	0	0	0	0	1	0	0	0	0	0	0	0
{Abe,Bob}	0	0	0	0	0	0	0	0	0	0	1	0
{Carl}	0	0	0	0	0	0	0	0	1	0	0	0
{Abe,Carl}	0	0	0	0	0	0	0	0	0	0	1	0
{Bob,Carl}	0	0	0	0	0	0	0	0	0	0	1	0
{A,B,C}	0	0	0	0	0	0	0	0	0	0	1	0
{Don}	0	0	0	0	0	0	0	0	0	0	1	0
{Abe,Don}	0	0	0	0	0	0	0	0	0	0	1	0
{Bob,Don}	0	0	0	0	0	0	0	0	0	0	1	0
{A,B,D}	0	0	0	0	0	0	0	0	0	0	1	0
{Carl,Don}	0	0	0	0	0	0	0	0	0	0	1	0
{A,C,D}	0	0	0	0	0	0	0	0	0	0	1	0
{B,C,D}	0	0	0	0	0	0	0	0	0	0	1	0
{A,B,C,D}	0	0	0	0	0	0	0	0	0	0	0	1

Fig. 7.3. Belief measure and plausibility measure for M of Fig. 7.2

The belief measure adds information to the extent that all evidence of subsets with an evidence attached is incorporated. Another idea is followed by the plausibility measure. One asks which sets have a non-empty intersection with some set with an evidence attached and determines the least upper bound of all these.

The plausibility measure collects those pieces of evidence that do *not* indicate trust against occurrence of the event or non-void parts of it. The belief as well as the plausibility measure more or less precisely determine their original body of evidence.

7.5 Proposition. Should the body of evidence be concentrated on singleton sets only, the belief and the plausibility measure will coincide.

Proof: That M is concentrated on arguments which are singleton sets means that $M = a{:}M$ with a the partial diagonal relation describing the atoms of the ordering Ω. For Ω and a one can prove $(\Omega \cap \overline{\Omega{:}\mathbb{T}}){:}a = a$ as the only other element less or equal to an atom, namely the least one, has been cut out via $\overline{\Omega}$. Then

$$\begin{aligned}(\Omega \cap \overline{\Omega{:}\mathbb{T}})^{\mathsf{T}}{:}\Omega{:}M &= (\Omega^{\mathsf{T}} \cap \mathbb{T}{:}\overline{\Omega}^{\mathsf{T}}){:}\Omega{:}a{:}M && M = a{:}M \text{ and transposing}\\ &= \Omega^{\mathsf{T}}{:}(\Omega \cap \overline{\Omega{:}\mathbb{T}}){:}a{:}M && \text{mask shifting}\\ &= \Omega^{\mathsf{T}}{:}a{:}M && \text{see above}\\ &= \Omega^{\mathsf{T}}{:}M && \text{again since } M = a{:}M \quad \square\end{aligned}$$

One should compare this result with the former one assuming m to be a mapping putting $m := \varepsilon{:}M$. One may also try to go in reverse direction, namely from a measure back to a body of evidence.

7.6 Definition. Let some measure μ be given and define strict subset containment $C := \overline{\mathbb{I}} \cap \Omega$. We introduce two basic probability assignments, namely

i) $A_\mu := \mathtt{lubR}_E(C^{\mathsf{T}}{:}\mu)$, its **purely additive part**,

ii) $J_\mu := \mu_1 \otimes (\mu \cap \overline{\mathtt{lubR}_E(C^{\mathsf{T}}{:}\mu)})$, its **jump part**. $\square$

As an example, the purely additive part A_μ of the μ of Fig. 6.1 would assign in line {Abe,Bob} the value {high,fair} only as $\mu(\{\text{Abe}\}) = \{\text{high,lazy}\}$ and $\mu(\{\text{Bob}\}) = \{\text{medium,fair}\}$. In excess to this, μ assigns {high,good}, and is, thus, not additive or Bayesian. We have for A_μ taken only what could have been computed already by summing up the values attached to *strictly* smaller subsets. In J_μ the excess of μ to A_μ is collected. In the procedure for J_μ all the values attached to atoms of the lattice will be saved as from an atom only one step down according to C is possible. The value for the least element is, however, the least element of $\mathcal{L}$. Multiplication with μ_1 serves the purpose that rows full of **0** 's be converted to rows with the least element 0_E attached as a value.

Now some arithmetic on these parts is possible, not least providing the insight that a measure decomposes into an additive part and a jump part.

7.7 Proposition. Given the present setting, we have

i) $A_\mu \oplus J_\mu = \mu$.

ii) $\mu_{\text{belief}}(J_\mu) = \mu$. $\square$

In the real-valued case, this result is not surprising at all as one may always decompose into a part continuous from the left and a jump part.

In view of these results it seems promising to investigate in which way also concepts such as commonality, consonance, necessity measures, focal sets, and cores may be found in the relational approach. This seems particularly interesting as also the concepts of De Morgan triples have been transferred to the componentfree relational side. We leave this to future research.

8 Concluding Remark

There exists a bulk of literature around the topic of Dempster-Shafer belief. It concentrates mostly on work with real numbers and their linear order and applies traditional free-hand mathematics. This makes it sometimes difficult to follow the basic ideas, not least as authors are all too often falling back to probability considerations.

We feel that the componentfree relational reformulation of this field and the important generalization accompanying it is a clarification — at least for the strictly growing community of those who do not fear to use relations. Proofs may now be supported via proof systems. The results of this paper have been formulated also in the relational language TituRel [8,9], for which some system support is available making it immediately operational. Not least has it provided computation and representation of the example matrices.

Acknowledgement. My thanks go to the unknown referees for their detailed comments which considerably improved the exposition.

References

1. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support. Volume 14 of Theory and Decision Library, Series D: System Theory, Knowledge Engineering and Problem Solving. Kluwer Academic Publishers (1994)
2. Schmidt, G., Ströhlein, T.: Relationen und Graphen. Mathematik für Informatiker. Springer-Verlag (1989) ISBN 3-540-50304-8, ISBN 0-387-50304-8.
3. Schmidt, G., Ströhlein, T.: Relations and Graphs — Discrete Mathematics for Computer Scientists. EATCS Monographs on Theoretical Computer Science. Springer-Verlag (1993) ISBN 3-540-56254-0, ISBN 0-387-56254-0.
4. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press (1976)
5. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Annals of Math. Statistics **38** (1967) 325–339
6. Choquet, G.: Theory of capacities. Annales de l'Institut Fourier **5** (1953) 131–295
7. Sugeno, M., ed.: Industrial Applications fo Fuzzy Control. North-Holland (1985)
8. Schmidt, G.: Relational Language. Technical Report 2003-05, Fakultät für Informatik, Universität der Bundeswehr München (2003) 101 pages, `http://homepage.mac.com/titurel/Papers/LanguageProposal.html`.
9. Schmidt, G.: The Relational Language TituRel: Revised Version (2005) In preparation; see `http://homepage.mac.com/titurel/TituRel/LanguageProposal2.pdf`.