Some Algebraic Aspects of Semantic Uncertainty and Cognitive Biases

Günther Wirsching
Katholische Universität Eichstätt-Ingolstadt
Mathematisch-Geographische Fakultät
85072 Eichstätt, Germany
Email: guenther.wirsching@ku.de

Robert Lorenz
Universität Augsburg
Institut für Informatik
86159 Augsburg, Germany
Email: robert.lorenz@informatik.uni-augsburg.de

***Abstract*—Based on the concept of representing semantic uncertainty by a *valuation map* which assigns to certain *semantic units* valuations taken from an abstract *valuation set*, we provide formal proofs that certain reasonable properties of the operations *decision* and *reinforcement* are related to algebraic structures on the valuation set. On the other hand, we infer from psychological research that in human behaviour reasonable operations are distorted by cognitive biases. In a final section, we connect our results to applications and propose a method for modeling the influence of certain cognitive biases.**

I. Introduction

When designing an artifical system for cognitive infocommunications [1], it is difficult to avoid semantic uncertainty. The primary goal of this article was to support the view that an appropriate dealing with semantic uncertainty is reflected in certain algebraic laws. This would lead to algebraic structures which can be considered as homomorphic images of processes performed by human beings when dealing with semantic uncertainty.

While working on this problem, we saw that we can *formally prove*, based on relational calculus, that certain apparently reasonable ways of dealing with semantic uncertainty are *essentially equivalent* to certain algebraic laws. On the other hand, we recognized that psychological research has revealed that *real* human behaviour is often dominated by *cognitive biases*, which prevents humans from behaving *apparently reasonable* [2]. Finally, we saw that there are effective methods for reconciling these two contradicting views.

In this paper, we focus attention on two psychological operations: *decision* and *reinforcement*. For making a mathematical analysis possible, it is necessary to fix precisely what we mean by these operations. Our modeling assumptions are as follows.

1) We call *decision* the selection of certain semantic units and neglecting others.
2) We call *reinforcement* any changes of valuations by incorporating new information provided by a second valuation map. In most cases, reinforcement will have the effect that the valuations of certain semantic units are 'strengthened', while the valuations of others are 'weakened'.

The plan of the paper is as follows. In section II, we start with a parsimonious concept *valuation map*, which is just an assignment of *valuations* to *semantic units*. Here we assume that the *meaning* of an utterance comes as set of clearly distinguishable *semantic units*, and a valuation map assigning to each semantic unit one or more numbers indicating confidence, probability, or, in general, some figures related to something like 'degree of reliability'. The valuation map concept allows us to relate certain psychological operations to set-theoretic and algebraic structures on the set of possible valuations. On the basis of these modeling assumptions, we exhibit relations between laws concerning semantic operations on one side, and algebraic laws concerning binary operations on the set of valuations on the other side. This is mainly original research, but related to known results in computer science [3]. The final section III relates our results to applications and discusses an idea how to incorporate certain cognitive biases.

II. Valuations and operations

An essential point in modeling semantic uncertainty is to associate *valuations* to *semantic units* which enable the system to perform the basic operations *decision* and *reinforcement*. Before diving into mathematical investigations, let us consider an example which in the sequel will be used to illustrate our results.

Example 1: Let us imagine a travel information system. Before giving information to the user, it needs some input from the user, given in one or more dialogue steps. Let us assume that the input is given as spoken utterances, and that the system contains a speech recognizer producing as recognition result an N-best list. Let us further assume that the system is able to associate to each entry of the N-best list a *meaning*, and that the meaning is represented by *feature-value* pairs, e. g.:

- Destination: Prague.
- Date for traveling: next Tuesday.
- Departure time: after 8 a.m.
- Arrival time: before midnight.

We consider each individual feature-value pair as a *semantic unit*. It may happen that different wordings convey the same semantic unit. For example, the meanings of the following phrases both contain the semantic unit [Destination : Prague]:

- I want to travel to Prague.
- The destination is Prague.

For simplicity, let us assume that each utterance conveys exactly one semantic unit. If it happens that the speech recognizer associates to different wordings the same semantic unit, we consider them as equivalent. Under these assumptions, a recognition result comes as an N-best list of semantic units with the property that a semantic unit occurs at most once in the N-best list. Moreover, let us assume that the system provides several data concerning reliability of the recognition results. For instance, such data may comprise acoustic scores and language model scores computed by the speech recognizer, data derived from the context of using this system, world knowledge like importance of certain destinations as touristic or business destination, and possibly data stemming from personalization. We shall refer to these data as the *set of scores*, and we assume that the *valuation* of a semantic unit is computed from the set of scores coming with the semantic unit. At this stage it is possible to insert an 'interface' distorting the computation in such a way that certain cognitive biases are modeled. For instance, in [4] a method is given which models the loss aversion bias.

Example 2: Let us return to the speech recognition example described above. In this case, for computing the valuation of a semantic unit, we should take into account acoustic scores and information about the context, e. g., about the preceding system prompt and the situational context.

What is the meaning of *decision* and *reinforcement* in this example? As a recognition result may consist of several alternatives, the system eventually has to *decide* which alternatives should be retained and which ones should be rejected. It is not assumed that the system always decides immediately after receiving a recognition result. If the system is not sure about the intention of the user, it may delay a decision to the future, and ask for repetition or further information. We call *reinforcement* the combination of recognition results from two or more utterances.

For fixing notation, assume that $\mathbb{U}$ is an arbitrary set of semantic units, and let $\mathbb{X}$ denote a for the moment unknown set of possible valuations. Then a *valuation map* is a map

$$\Phi : \mathbb{U} \to \mathbb{X}. \tag{1}$$

A map with target $\mathbb{X}$ which is only defined on a subset $\mathbb{V} \subseteq \mathbb{U}$ is called a *partial valuation map*. The domain of definition of a (partial) valuation map is denoted by $\mathbb{D}(\Phi)$. This concept is similar to the 'valuation function' considered in [5, p. 4]. In this setting, *decision* is an operation restricting a given (partial) valuation map Φ to some subset

$$\mathbb{D}(\nabla\Phi) \subseteq \mathbb{D}(\Phi) \subseteq \mathbb{U},$$

leading to a partial valuation $\nabla\Phi := \Phi|_{\mathbb{D}(\nabla\Phi)}$ which only valuates the semantic units corresponding to the decision. The elements of $\mathbb{D}(\nabla\Phi)$ are called *winners* of the decision. For instance, if $\mathbb{U} = \{u_1, u_2\}$, and the system decides to retain only the semantic unit u_1, then u_1 is the only winner of the decision, and $\nabla\Phi = \Phi|_{\{u_1\}}$. Note that we do not exclude that $\mathbb{D}(\nabla\Phi) = \varnothing$; but we say that a decision operation ∇ is *non-vanishing*, if, for any valuation map Φ, any non-empty subset of $\mathbb{U}$ contains at least one winner. Similarly, *reinforcement* is an operation taking two valuation maps Φ_1 and Φ_2 as input and producing a new valuation map $\Phi_1 \,\&\, \Phi_2$ with the property

$$\mathbb{D}(\Phi_1 \,\&\, \Phi_2) = \mathbb{D}(\Phi_1) \cap \mathbb{D}(\Phi_2).$$

The idea is to define appropriate mathematical structures on the set $\mathbb{X}$ from which these operations can be derived. The condition is that the valuations of the outcome should only depend on the valuations of semantic units, and not on their semantic content. In [6], an operation is called *consistent* w.r.t. valuation, if it fulfills this condition.

A. Consistent decision operations and binary relations

In case of the decision operation, the system basically has to decide in favor of one of two semantic units. If $\Phi : \mathbb{U} \to \mathbb{X}$ is a valuation map, *consistency* of decision with valuation means that $\mathbb{D}(\nabla\Phi)$ is a union of complete level sets of Φ, i.e., given a level $c \in \mathbb{X}$, then

$$\Phi^{-1}(c) \subseteq \mathbb{D}(\nabla\Phi) \quad \text{or} \quad \Phi^{-1}(c) \cap \mathbb{D}(\nabla\Phi) = \varnothing. \tag{2}$$

The basic mathematical concept used for describing the concept 'decision between two semantic units' is that of a *binary relation* on a set $\mathbb{X}$, formally defined as a *set of pairs*, or as a subset of the cartesian product $\mathbb{X}\times\mathbb{X}$. If $\rhd \subseteq \mathbb{X}\times\mathbb{X}$ denotes a binary relation, then an expression '$x \rhd y$' just means that the pair $(x, y) \in \rhd$. For defining a decision operation out of a binary relation, we extend this notation to arbitrary subset $\mathbb{Y} \subseteq \mathbb{X}$ by setting

$$x \rhd \mathbb{Y} \quad :\Leftrightarrow \quad \forall y \in \mathbb{Y} : x \rhd y. \tag{3}$$

In case $x \rhd \mathbb{Y}$ we say that x *dominates* $\mathbb{Y}$ *w.r.t.* $\rhd$. Given a valuation map $\Phi : \mathbb{U} \to \mathbb{X}$, and a subset $\mathbb{V} \subseteq \mathbb{U}$, we say that an element $u \in \mathbb{U}$ *(Φ-)dominates* $\mathbb{V}$ *(w.r.t. B)*, if $\Phi(u) \, B \, \Phi(\mathbb{V})$. For instance, for any valuation map Φ, any element $u \in \mathbb{U}$ Φ-dominates the empty set $\varnothing$ (w.r.t. any binary relation).

Theorem 1: Let $\mathbb{U}$ be an arbitrary set of semantic units, and let $\rhd$ denote a binary relation on a set $\mathbb{X}$. Given a valuation maps $\Phi : \mathbb{U} \to \mathbb{X}$, setting

$$u \in \mathbb{D}(\nabla\Phi) \;:\Leftrightarrow\; \Phi(u) \rhd \Phi(\mathbb{U}) \tag{4}$$

defines a decision operation ∇ on valuation maps which is consistent with valuation.

Proof: In order to prove consistency, it suffices to prove assertion (2) from (4). But this is almost immediate: If $\Phi : \mathbb{U} \to \mathbb{X}$ is a valuation map, and if $u_1, u_2 \in \mathbb{U}$ are two semantic units with equal valuation, $\Phi(u_1) = \Phi(u_2)$, then condition (4) shows that

$$u_1 \in \mathbb{D}(\nabla\Phi) \;\Leftrightarrow\; u_2 \in \mathbb{D}(\nabla\Phi),$$

which reformulates (2). ■

Example 3: Let us return to our travel information system example with recognition result given in an N-best list of semantic units, and suppose that we receive the list

1) [Date : tomorrow] ("I want to travel tomorrow"),
2) [Dest : Moscow] ("I want to travel to Moscow"),

where each item is equipped with a set of scores. Assuming that the *valuation* of a semantic unit is computed from its set of scores, *consistency* of decision with valuation means that decision is only based on valuation. In our recognition result, acoustic scores are likely to be similar–assume they are equal. If valuation is based only on acoustic scores, a decision which is consistent with valuation must either take both alternatives, or reject both. If we had also scores from the context, e. g., if the system expects as input the date of traveling, we could incorporate these context scores into valuation. In this case, a decision for the first recognition alternative while rejecting the second can be consistent with valuation.

B. Non-vanishing and effective decisions

The construction in (4) does not exclude the case $\mathbb{D}(\nabla\Phi) = \varnothing$. As an example, consider the three-element set $\mathbb{X} = \{x, y, z\}$ and the binary relation

$$\rhd := \{(x, y), (y, z), (z, x)\} \subseteq \mathbb{X} \times \mathbb{X},$$

and the valuation map $\Phi : \{u_1, u_2, u_3\} \to \mathbb{X}$ given by $\Phi(u_1) = x$, $\Phi(u_2) = y$, and $\Phi(u_3) = z$. In this situation, there is no semantic unit whose valuation dominates all the others, whence $\mathbb{D}(\nabla\Phi) = \varnothing$.

In mathematics, there are some coined verbal expressions referring to certain properties of binary relations. A binary relation $B \subset \mathbb{X} \times \mathbb{X}$ is called

reflexive,	if $\forall x \in \mathbb{X} : x\,B\,x$,
total,	if $\forall x, y \in \mathbb{X} :\ x\,B\,y$ or $y\,B\,x$,
antisymmetric,	if $x\,B\,y$ and $y\,B\,x$ imply $x = y$,
transitive,	if $x\,B\,y$ and $y\,B\,x$ imply $x = y$.

Note that any total binary relation is also reflexive.

For having a property of binary relations ensuring that the induced decision operation is non-vanishing, it appears natural to call a binary relation *subset-topped*, if any non-empty subset of $\mathbb{X}$ contains an element dominating the subset via that binary relation. Unfortunately, the usual order relation $\geqslant$ on the set $\mathbb{R}$ of real numbers is not subset-topped in this sense, as, for instance, an open interval doesn't contain a dominating element.

In order to come to a more appropriate notion, let c denote an arbitrary cardinal number, finite or transfinite. A binary relation is called *c-subset-topped,* if any set of cardinality $\leqslant c$ contains a dominating element, and it is called *finite-subset-topped,* if it is n-subset-topped for any finite cardinal n. For two cardinals $c \geqslant d > 0$, any c-subset-topped binary relation is also d-subset-topped. A binary relation is reflexive, if and only if it is 1-subset-topped. And a binary relation is total, if and only if it is 2-subset-topped. In addition, any total, transitive relation $\rhd$ is finite-subset-topped, as arranging the elements of a finite subset in a chain

$$x_1 \rhd x_2 \rhd \ldots \rhd x_n$$

implies by transitivity that $x_1 \rhd \{x_1, \ldots, x_n\}$.

Denoting the cardinality of a given set $\mathbb{U}$ by $|\mathbb{U}|$, we have the following characterization of non-vanishing decision operations:

Theorem 2: Let $\mathbb{U}$ be a non-empty set of semantic units, let $\rhd$ be a binary relation on $\mathbb{X}$, and let ∇ be defined according to (4). Then ∇ is non-vanishing, if and only if $\rhd$ is $|\mathbb{U}|$-subset-topped.

Proof: 'Only if'-part: Assume the ∇ is non-vanishing, and let $\mathbb{Y} \subseteq \mathbb{X}$ be a subset satisfying $0 < |\mathbb{Y}| \leqslant |\mathbb{U}|$. Then there is a valuation map $\Phi : \mathbb{U} \to \mathbb{X}$ such that $\Phi(\mathbb{U}) = \mathbb{Y}$. Because ∇ is non-vanishing, we infer that $\mathbb{D}(\nabla\Phi) \neq \varnothing$, which implies by (4) that there is an element $u \in \mathbb{D}(\nabla\Phi)$ with the property $x = \Phi(u) \rhd \Phi(\mathbb{U}) = \mathbb{Y}$. This proves that $\rhd$ is $|\mathbb{U}|$-subset-topped.

'If'-part: Let $\rhd$ be $|\mathbb{U}|$-subset-topped, and let $\Phi : \mathbb{U} \to \mathbb{X}$ be a valuation map. Because $|\Phi(\mathbb{U})| \leqslant |\mathbb{U}|$, there exists $x \in \Phi(\mathbb{U})$ such that $x \rhd \Phi(\mathbb{U})$. Choose $u \in \mathbb{U}$ satisfying $\Phi(u) = x$. According to (4), $u \in \mathbb{D}(\nabla\Phi)$, which proves $\mathbb{D}(\nabla\Phi) \neq \varnothing$. ■

Let us call a decision operation ∇ *effective*, if for any two semantic units $u_1, u_2 \in \mathbb{U}$ with different valuation $\Phi(u_1) \neq \Phi(u_2)$, at most one of them is a winner. In other words, ∇ is *effective* if its result $\mathbb{D}(\nabla\Phi)$ does not contain two semantic units with different valuations. We have the following characterization of effective decision operations:

Theorem 3: Let $\mathbb{U}$ be a set of semantic units, let $\rhd$ be a binary operation, and let ∇ be the decision operator induced by $\rhd$ via (4). Then ∇ is effective, if and only if $\rhd$ is antisymmetric.

Proof: 'Only if'-part: Let $x, y \in \mathbb{X}$ with $x \rhd y$ and $y \rhd x$. Let further Φ be a valuation map with $\mathbb{D}(\Phi) = \{u_1, u_2\}$, $\Phi(u_1) = x$ and $\Phi(u_2) = y$. By (4) we get $\mathbb{D}(\nabla\Phi) = \{u_1, u_2\}$. If ∇ is effective, then $x = \Phi(u_1) = \Phi(u_2) = y$, i.e. $\rhd$ is antisymmetric.

'If'-part: Let Φ be a valuation map and let $u_1, u_2 \in \mathbb{D}(\nabla\Phi)$. By (4) we get $\Phi(u_1) \rhd \Phi(u_2)$ and $\Phi(u_2) \rhd \Phi(u_1)$. If $\rhd$ is antisymmetric, then $\Phi(u_1) = \Phi(u_2)$, which means that ∇ is effective. ■

Example 4: If a recognition result consists of two recognition alternatives with different valuations, then a *non-vanishing* decision operation takes at least one them, while an *effective* decision operation takes at most one of them.

C. Stability of a decision operation

Let $\rhd$ denote a binary operation. The *symmetric part* of $\rhd$ is defined by

$$x \sim y \quad :\Leftrightarrow \quad x \rhd y \,\wedge\, y \rhd x. \tag{5}$$

A $\rhd$*-clique* is a subset $\mathbb{Y} \subseteq \mathbb{X}$ satisfying $\forall x, y \in \mathbb{Y} : x \sim y$ which is maximal w.r.t. set inclusion. The binary relation $\rhd$ is called *clique-decomposable*, if $\mathbb{X}$ can be decomposed into a union of pairwise disjoint $\rhd$-cliques, or, equivalently, if its symmetric part $\sim$ is reflexive and transitive. Note that any reflexive, antisymmetric binary relation is clique-decomposable, because reflexivity implies that any element belongs to a clique, and antisymmetry implies that any clique consists of at most one element.

Moreover, given a valuation map $\Phi : \mathbb{U} \to \mathbb{X}$ and a binary relation $\rhd$ on $\mathbb{X}$, two semantic units $u_1, u_2 \in \mathbb{U}$ are called Φ*-equivalent*, if $\Phi(u_1) \rhd \Phi(u_2)$ and $\Phi(u_2) \rhd \Phi(u_1)$; formally

$$u_1 \sim_\Phi u_2 \quad :\Leftrightarrow \quad \Phi(u_1) \sim \Phi(u_2). \tag{6}$$

We call a decision operation ∇ *equivalent value stable*, if, for any valuation map $\Phi : \mathbb{U} \to \mathbb{X}$, the binary relation $\sim_\Phi$ is reflexive and transitive. Note that, if decision is equivalent value stable, then both binary relations $\rhd$ and its symmetric part $\sim$ are reflexive.

Theorem 4: Let $\mathbb{U}$ be a non-empty set of semantic units, let $\rhd$ be a $|\mathbb{U}|$-subset-topped relation on a set $\mathbb{X}$. Then the decision operation ∇ on valuation maps induced by $\rhd$ is equivalent value stable, if and only if $\rhd$ is clique-decomposable.

Proof: As $\mathbb{U} \neq \varnothing$, we have $|\mathbb{U}| \geqslant 1$, from which we infer that $\rhd$ is reflexive.

'Only if'-part: Assume the ∇ is equivalent value stable. For proving that $\rhd$ is clique-decomposable, it remains to show that $\sim$ is transitive. Let $x, y, z \in \mathbb{X}$ such that $x \sim y$ and $y \sim z$. Further let Φ be a valuation map with $\mathbb{D}(\Phi) = \{u_1, u_2, u_3\}$, $\Phi(u_1) = x$, $\Phi(u_2) = y$, and $\Phi(u_3) = z$. Then $u_1 \sim_\Phi u_2 \sim_\Phi u_3$. If ∇ is equivalent value stable, then $u_1 \sim_\Phi u_3$, and (6) implies $x \sim z$, whence $\sim$ is tansitive.

'If'-part: Let Φ be a valuation map. As reflexivity of $\sim$ implies that the binary relation $\sim_\Phi$ on $\mathbb{U}$ is also reflexive, we only have to prove that $\sim_\Phi$ is transitive. To this end, let $u_1, u_2, u_3 \in \mathbb{D}(\Phi)$ such that $u_1 \sim_\Phi u_2$ as well as $u_2 \sim_\Phi u_3$. Then (6) gives $\Phi(u_1) \sim \Phi(u_2) \sim \Phi(u_3)$. If $\rhd$ is clique-decomposable, then $\Phi(u_1) \sim \Phi(u_3)$, which means that $u_1 \sim_\Phi u_3$, whence $\sim_\Phi$ is transitive. ■

Let us call a decision ∇ *subset-stable*, if, for any valuation map $\Phi : \mathbb{U} \to \mathbb{X}$ and any subset $\mathbb{V} \subseteq \mathbb{U}$, its result doesn't change when it is performed first on $\mathbb{V}$ and then on the complement $\mathbb{U} \setminus \mathbb{V}$ joined to the winners $\mathbb{W} := \mathbb{D}(\nabla(\Phi|_\mathbb{V}))$ of decision on the subset. Using the notation

$$\mathbb{V}' := (\mathbb{U} \setminus \mathbb{V}) \cup \mathbb{W} = (\mathbb{U} \setminus \mathbb{V}) \cup \mathbb{D}(\nabla(\Phi|_\mathbb{V})), \tag{7}$$

decision is called *subset-stable* if, for any subset $\mathbb{V} \subseteq \mathbb{U}$,

$$\mathbb{D}\left(\nabla\left(\Phi|_{\mathbb{V}'}\right)\right) = \mathbb{D}(\nabla\Phi). \tag{8}$$

Theorem 5: Let $\mathbb{U}$ be a set of semantic units satisfying $|\mathbb{U}| \geqslant 3$, and let $\rhd$ be a $|\mathbb{U}|$-subset-topped relation on a set $\mathbb{X}$ which is also clique-decomposable. Then the decision operation ∇ on valuation maps induced by $\rhd$ is subset-stable, if and only if $\rhd$ is transitive.

Proof: As $\rhd$ is clique-decomposable, it follows from that its symmetric part $\sim$ is reflexive and transitive. Formally, we have to prove that

$$\nabla \text{ is subset-stable} \iff \rhd \text{ is transitive.} \tag{9}$$

'$\Rightarrow$': Let $x \rhd y \rhd z$ be a chain in $\mathbb{X}$, then we have to prove that $x \rhd z$.

Because $\mathbb{U}$ has at least three elements, we can choose semantic units $u_1, u_2, u_3 \in \mathbb{U}$, and a (partial) valuation map $\Phi : \{u_1, u_2, u_3\} \to \mathbb{X}$ with $\Phi(u_1) = x$, $\Phi(u_2) = y$, and $\Phi(u_3) = z$. We consider the following four cases:

First case: $y \not\rhd x$ and $z \not\rhd y$. If $x \not\rhd z$, the set $\mathbb{Y} := \{x, y, z\}$ wouldn't have a dominating element, whence $\rhd$ couldn't be 3-subset-topped. Therefore, $x \rhd z$.

Second case: $y \not\rhd x$ and $z \rhd y$. Assume $x \not\rhd z$, and fix the subset $\mathbb{V} := \{u_1, u_3\} \subseteq \mathbb{U}$. By (4) we infer $\mathbb{D}(\nabla(\Phi|_\mathbb{V})) = \{u_3\}$, hence $\mathbb{V}' = \{u_2, u_3\}$ according to (7). As $y \rhd z$ and $z \rhd y$, we deduce $\mathbb{D}(\nabla(\Phi|_{\mathbb{V}'})) = \{u_2, u_3\}$. On the other hand, $y \not\rhd x$ implies $\mathbb{D}(\nabla\Phi) = \{u_3\}$, contradicting that ∇ is subset-stable. Hence $x \rhd z$.

Third case: $y \rhd x$ and $z \not\rhd y$. Fix $\mathbb{V} = \{u_2, u_3\}$. By (4) we get $\{u_2\} = \mathbb{D}(\nabla(\Phi|_\mathbb{V}))$, and (7) gives $\mathbb{V}' = \{u_1, u_2\}$. We deduce $\mathbb{D}(\nabla(\Phi|_{\mathbb{V}'})) = \{u_1, u_2\}$. If ∇ is subset-stable, then $\mathbb{D}(\nabla\Phi) = \{u_1, u_2\}$, implying $x \rhd z$.

Fourth case: $y \rhd x$ and $z \rhd y$. In this case, $u_1 \sim_\Phi u_2 \sim_\Phi u_3$. Since ∇ is equivalent value stable, this implies $u_1 \sim_\Phi u_3$, whence $x \sim z$, which implies $x \rhd z$.

'$\Leftarrow$' of (9): Let $\Phi : \mathbb{U} \to \mathbb{X}$ be a valuation map, let $\mathbb{V} \subseteq \mathbb{U}$, and use the notation

$$\mathbb{V}' = (\mathbb{U} \setminus \mathbb{V}) \cup \mathbb{D}(\nabla(\Phi|_\mathbb{V})$$

as in (7). We can assume without loss of generality that $\mathbb{V} \setminus \mathbb{D}(\nabla(\Phi|_\mathbb{V})) \neq \varnothing$, since otherwise $\mathbb{V}' = \mathbb{U}$, and (8) would become trivial. Moreover, since $\rhd$ is $|\mathbb{U}|$-subset-topped, ∇ is non-vanishing, which means $\mathbb{D}(\nabla(\Phi|_\mathbb{V})) \neq \varnothing$.

It remains to prove the set equation $\mathbb{D}(\nabla\Phi) = \mathbb{D}(\nabla(\Phi|_\mathbb{W}))$. For proving '$\subseteq$', let $u \in \mathbb{D}(\nabla\Phi)$. Then $\Phi(u) \rhd \Phi(\mathbb{U})$, from which we get $\Phi(u) \rhd \Phi(\mathbb{V}')$. Since u cannot be an element of $\mathbb{V} \setminus \mathbb{D}(\nabla(\Phi|_\mathbb{V}))$, we infer $u \in \mathbb{D}(\nabla(\Phi|_{\mathbb{V}'}))$.

For proving '$\supseteq$', let $u \in \mathbb{D}(\nabla(\Phi|_{\mathbb{V}'}))$. If $u \in \mathbb{D}(\nabla(\Phi|_\mathbb{V}))$, then u dominates $\mathbb{V}$ and $\mathbb{U} \setminus \mathbb{V}$, i.e., $u \in \mathbb{D}(\nabla\Phi)$. If $u \in \mathbb{U} \setminus \mathbb{V}$, then, in particular, u dominates $\mathbb{D}(\nabla(\Phi|_\mathbb{V}))$. Finally, let $w \in \mathbb{V} \setminus \mathbb{D}(\nabla(\Phi|_\mathbb{V}))$. Choose $v \in \mathbb{D}(\nabla(\Phi|_\mathbb{V}))$. Then $u \rhd v \rhd w$. If $\rhd$ is transitive, then $u \rhd w$. Therefore, u dominates the whole set $\mathbb{U}$, i.e., $u \in \mathbb{D}(\nabla\Phi)$. ■

By theorem 2, we know that whenever a binary relation $\rhd$ on the set of valuations induces a non-vanishing decision operation, then $\rhd$ has to be $|\mathbb{U}|$-subset-topped. If $\mathbb{U} \neq \varnothing$, this implies that $\rhd$ is reflexive. In case $\rhd$ is clique-decomposible, theorem 5 enables us to extend this result:

- If the induced decision operation is subset-stable, then $\rhd$ has to be transitive, additionally. This is related to the known concept *preorder*, where, by definition, a *preorder* is a binary relation which is both reflexive and transitive. Observe a total preorder is always finite-subset-topped.
- If the induced decision operation is effective and subset-stable, then $\rhd$ has to be antisymmetric and transitive, additionally. Here we are in the well-studied realm of *total orders*.

D. The intuitive notion 'higher'

A mathematical structure giving a notion of 'higher' is, for instance, the structure called *totally ordered set*. For defining this in mathematical terms, it is possible to do this either with the concept *total strict order*, or with the concept *total (non-strict) order*, where the latter is more common but less intuitive.

For serving intuition, it is worth to define the *irreflexive part* B^* of a binary relation B by

$$x\, B^*\, y \quad :\Leftrightarrow \quad x\, B\, y \quad \text{and} \quad x \neq y.$$

A binary relation B^* is called *trichotomic*, if, for arbitrary $x, y \in \mathbb{X}$, exactly one of the following assertions is true:

$$x\, B^*\, y, \quad x = y, \quad y\, B^*\, x. \tag{10}$$

Note that a total binary relation is antisymmetric, if and only if its irreflexive part is trichotomic.

In mathematics and computer science, it has turned out that a reflexive extension 'higher or equal', or, equivalently, 'not smaller', is also very useful. Formally, a binary relation is called a *strict total order*, if it is transitive and trichotomic. On the other hand, a *(non-strict) total order* is a binary relation which is total, reflexive, antisymmetric, and transitive.

The psychological content of the verbal expression 'higher' is captured in the mathematical notions *trichotomy* and *transitivity*. By the remark following (10), a non-strict total order is nothing but a reflexive binary relation whose irreflexive part captures the psychological content of a verbal expression like 'higher' or 'smaller'.

Example 5: In some situations, human judgements like 'a longer line' or 'a shorter line' do not coincide with physical measurement. A famous example is an optical illusion called *Müller-Lyer illusion*, where two lines of equal length are presented in different graphical contexts, which misleads virtually all human beings who see the pictures for the first time. If *decision* is interpreted as *decision for the longer line*, then the context-dependency shows that this decision operation is not subset-stable. By theorem 5 we infer that such a decision cannot be described by a total order, in perfect harmony with experimental data.

E. Algebraic structures

It is possible to connect the properties of binary relations investigated in the previous sections to properties of binary operations, which are commonly used for defining algebraic structures. To fix ideas, a *finitary operation* on a given set $\mathbb{S}$ is a map f taking a finite number of input elements $x_1, \ldots, x_n \in \mathbb{S}$ to produce an output $f(x_1, \ldots, x_n) = y \in \mathbb{S}$. The most common types of finitary operations are *constants* $c \in \mathbb{S}$, considered as 'zero-ary operations', i.e., maps with zero input elements, and *binary operations*, considered as maps

$$\star : \mathbb{S} \times \mathbb{S} \to \mathbb{S}, \qquad (x, y) \mapsto x \star y \in \mathbb{S}. \tag{11}$$

An *algebraic structure* consists of a set $\mathbb{S}$, called *carrier set* or *underlying set*, and some given finitary operations, subject to a number of logical assertions called *axioms* of the algebraic structure. Widely used axioms for binary operations are, for instance, the *law of associativity*,

$$\forall x, y, z \in \mathbb{X} : \quad x \star (y \star z) = (x \star y) \star z \,,$$

and the *law of commutativity*,

$$\forall x, y \in \mathbb{X} : \quad x \star y = y \star x \,.$$

An important axiom for the connection of binary relations and operations is the *law of idempotency*:

$$\forall x \in \mathbb{X} : \quad x \star x = x \,.$$

We found that a stronger axiom, which we call the *law of inertness*,

$$\forall x, y \in \mathbb{X} : \quad x \star y \in \{x, y\} \,, \tag{12}$$

is necessary. Our result is as follows.

Theorem 6: Let $\mathbb{X}$ be an arbitrary set.

1) If $\star$ is an inert binary operation on $\mathbb{X}$, then the *associated successor relation* $\succcurlyeq_\star$ on $\mathbb{X}$ defined by $x \succcurlyeq_\star y :\Leftrightarrow \big(x = x \star y \vee x = y \star x\big)$ is total.
2) If $\succcurlyeq$ is a total binary relation on $\mathbb{X}$, then the *associated maximum operation* $\star_\succcurlyeq$ defined by $x \star_\succcurlyeq y := x$, if $x \succcurlyeq y$, and $x \star_\succcurlyeq y := y$ otherwise, is an inert binary operation on $\mathbb{X}$.
3) A total binary relation is antisymmetric, if and only if its associated maximum operation is commutative.
4) A total binary relation is transitive, if and only if its associated maximum operation is associative.

Due to lack of space, we omit a formal proof of this theorem. Similar results can be found in the literature: In [3, Chap. 4], it is shown that the first binary operation in a *dioid* (an algebraic structure with two binary operations where the first one is assumed to be idempotent, associative, and commutative) gives rise to the construction of a partial order. Lemma 4.30 in [3] proves that this partial order is total, if and only if the first binary operation has the property given in (12).

The consequence of theorem 6 is that, whenever a binary relation $\rhd$ on the set of valuations induces a decision operation which is consistent with valuation, effective, and subset-stable, then $\rhd$ is total, and the associated binary operation $\star_\rhd$ is inert, commutative, and associative.

Example 6: Taking up again our speech recognition example, it appears reasonable to postulate that any two possible valuations can be compared. Therefore, it is quite natural to require that the first binary operation of the algebraic structure on the set of possible valuations is inert.

F. The reinforcement operation

As mentioned before, there is another psychologically important operation on valuation maps: *reinforcement*. Let $\Phi_1 : \mathbb{U} \to \mathbb{X}$ be a given valuation map. When new information is received, it should be incorporated into the information represented by Φ_1. Technically, suppose that the new information is represented as another valuation map Φ_2, then this incorporation produces a new valuation map $\Phi_2 \,\&\, \Phi_2$. It appears appropriate to call this operation *reinforcement*, as any new information should make the valuation more precise—despite the fact that reinforcement may strengthen some semantic units while weakening others.

According to [6], an operation is called *consistent with valuation*, if, for each individual semantic unit $u \in \mathbb{U}$, its outcome valuation $(\Phi_1 \,\&\, \Phi_2)(u)$ only depends on the incoming valuations $\Phi_1(u), \Phi_2(u) \in \mathbb{X}$, and not directly on the involved semantic contents. Therefore, reinforcement is consistent with valuation, if and only if it is induced by a binary operation $\circ$ on $\mathbb{X}$, formally:

$$\Phi_1 \,\&\, \Phi_2 : \mathbb{U} \to \mathbb{X}, \quad (\Phi_1 \,\&\, \Phi_2)(u) := \Phi_1(u) \circ \Phi_2(u). \tag{13}$$

Example 7: In our speech recognition example, reinforcement comes in when a second utterance is recognized. Each recognition result gives a valuation map on the set of semantic units. Given the first valuation map, we use the binary operation *reinforcement* for incorporating the information from the second recognition result.

The next step is to ponder on algebraic laws concerning the binary operation $\circ$ on the set of possible valuations. Having

(13) as background, it is not reasonable to postulate that $\circ$ should be inert—in this respect, *reinforcement* is different from *decision*.

The link between *commutativity* of the binary operation $\circ$ inducing a reinforcement operation and psychology is that commutativity of $\circ$ is equivalent to *stability* of reinforcement w.r.t. ordering of the presented information. But there is an important and well-studied psychological effect called *priming* [2], which means that previous information usually influences updating information, even if the previous information is received subconsciously. Therefore, in human psychology, reinforcement is clearly not stable w.r.t. ordering of presented information. Hence, for designing an aritifical system whose behavior should be agreeable for human beings, commutativity of the second binary operation on $\mathbb{X}$ is not an essential postulate.

The law of *associativity* of the binary operation $\circ$ is equivalent to *stability* of reinforcement w.r.t. aggregation of information: associativity of $\circ$ means that it is unimportant for valuation whether new information comes piecewise in two chunks, or combined in one chunk. A psychological effect possibly contradicting this kind of stability of reinforcement is the *halo effect* [2].

A further mathematical notion not mentioned before is *distributivity*: given two binary operations $\star$ and $\circ$ on $\mathbb{X}$, the operation $\circ$ is said to *distribute over* $\star$, provided

$$\forall x, y, z \in \mathbb{X} : \quad \begin{aligned} x \circ (y \star z) &= (x \circ y) \star (x \circ z), \\ (x \star y) \circ z &= (x \circ z) \star (y \circ z). \end{aligned} \tag{14}$$

This is the non-commutative formulation; if $\circ$ were assumed to be commutative, it would suffice to postulate only one of the two equations. The psychological pendant to distributivity is *stability* of decision w.r.t. reinforcement. More precisely, the first equation in (14) means that if two semantic units are equally valuated by a given valuation map Φ_1, and a new valuation map Φ_2 comes in, then it is unimportant whether decision is made on the basis of Φ_2, and then reinforced, or reinforcement is made first and then decision. The second equation in (14) means that if the incoming new valuation map puts the equal valuations on two semantic units u_1 and u_2, then it is unimportant whether decision comes first and then reinforcement, or reinforcement comes first and then decision. It is obvious that real human behaviour often violates this—in fact, real life decisions strongly depend on availability of information.

III. Applications

An overview of results achieved in section II is given in Table I: the first column refers to mathematical laws, the second one to semantic operations, and the third column indicates a cognitive bias possibly violating the law. As discussed in [6], the algebraic laws which we related to psychology lead to the algebraic structure *semiring*, which opens the use of weighted finite state transducers as described in [7] even for processing semantics. In addition, semirings are the basic algebraic strcuture used for weighted Petri net transducers described in [8].

It remains the problem how to deal with cognitive biases. We've seen in [6] that priming can be modeled by using a non-

TABLE I. Algebraic laws vs. cognitive biases

Mathematical operation	**Semantic operation**	**Cognitive bias**
First binary operation $\star$	Decision	
– inert	– non-vanishing	
– commutative	– effective	Müller-Lyer paradox
– subset-stable	– associative	
Second binary operation $\circ$	Reinforcement	
– well-defined	– consistent	
– commutative	– order-stable	Priming
– associative	– aggregation-stable	Halo effect
$\circ$ distributes over $\star$	Decision is stable w.r.t. reinforcement	Context-dependency of decisions

commutative semiring which is constructed there. The reason for cognitive biases like anchoring or the halo effect seems to be that human behaviour is more context-dependent than mathematical operations. An idea to incorporate more context-sensitivity is *spreading*. We show by an example that a natural spreading strategy provides a connection between rigorous algebraic laws and a softer dealing with semantic uncertainty which is able to model certain cognitive biases.

Example 8: Suppose that the set $\mathbb{U}$ of semantic units is equipped with *similarity weights* $\sigma : \mathbb{U} \times \mathbb{U} \to [0, 1]$, and consider valuation maps $\Phi : \mathbb{U} \to [0, 1]$ with the binary operations $\star = \max$ and $\circ$ meaning multiplication. Then *spreading* Σ operates on a valuation map Φ as follows:

$$\Sigma\Phi : \mathbb{U} \to \mathbb{R}, \quad \Sigma\Phi(u) := \max\{\sigma(u, x)\, \Phi(x) : x \in \mathbb{U}\}\,.$$

Under the natural requirement that any semantic unit $u \in \mathbb{U}$ is 'maximal self-similar', meaning $\sigma(u, u) = 1$, we get $\Sigma\Phi \geqslant \Phi$, which can influence decision. If we combine spreading and reinforcement, we get

$$\forall u \in \mathbb{U} : \quad \Phi_2(u) \circ \Sigma\Phi_1(u) \geqslant \Phi_2(u) \circ \Phi_1(u)\,,$$

and this can have an even stronger influence on decision. It follows that spreading can be used for stressing 'similar' semantic units. This can serve as a model for anchoring, or a halo effect.

References

[1] P. Baranyi and A. Csapó, "Definition and Synergies of Cognitive Infocommunications," *Acta Polytechnica Hungarica*, vol. 9, pp. 67–83, 2012.

[2] D. Kahneman, *Schnelles Denken, langsames Denken*. Siedler, 2012.

[3] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat, "Synchronization and linearity : an algebra for discrete event systems. Wiley Series in probability and mathematical statistics," 1992.

[4] P. Földesi and J. Botzheim, "Computational method for corrective mechanism of cognitive decision-making biases," in *2012 IEEE 3rd International Conference on Cognitive InfoCommunication, Košice, Slovakia, December 2-5, 2012*, 2012, pp. 211–215.

[5] N. H. Anderson, *Methods of Information Integration Theory*. New York: Academic Press, 1982.

[6] G. Wirsching, "Calculating semantic uncertainty," in *2012 IEEE 3rd International Conference on Cognitive InfoCommunication, Košice, Slovakia, December 2-5, 2012*, 2012, pp. 71–76.

[7] M. Mohri, "Weighted automata algorithms," in *Handbook of Weighted Automata*, pp. 213–254.

[8] R. Lorenz, M. Huber, and G. Wirsching, "On weighted petri net transducers," in *Application and Theory of Petri Nets and Concurrency - 35th International Conference, PETRI NETS 2014, Tunis, Tunisia, June 23-27, 2014. Proceedings*, ser. Lecture Notes in Computer Science, G. Ciardo and E. Kindler, Eds., vol. 8489. Springer, 2014, pp. 233–252.