The Context Model: An Integrating View of Vagueness and Uncertainty

Jörg Gebhardt and Rudolf Kruse
Technical University of Braunschweig, Germany

ABSTRACT

The problem of handling vagueness and uncertainty as two different kinds of partial ignorance has become a major issue in several fields of scientific research. Currently the most popular approaches are Bayes theory, Shafer's evidence theory, the transferable belief model, and the possibility theory with its relationships to fuzzy sets.

Since the justification of some of the mentioned theories has not been clarified for a long time, some criticism on these models is still pending. For that reason we have developed a model of vagueness and uncertainty—called the context model—that provides a formal environment for the comparison and semantic foundation of the referenced theories.

In this paper we restrict ourselves to the presentation of basic ideas keyed to the interpretation of Bayes theory and the Dempster–Shafer theory within the context model. Furthermore the context model is applied to show a direct comparison of these two approaches based on the well-known spoiled sandwich effect, the three prisoners problem, and the unreliable alarm paradigm.

KEYWORDS: ***Context model, Bayes theory, Dempster–Shafer theory, transferable belief model, decision-making, spoiled sandwich effect, three prisoners problem, unreliable alarm paradigm***

1. INTRODUCTION

The context model aims to provide integrating structures and concepts for well-founded handling of imperfect data, where a measure–theoretical

Address correspondence to Rudolf Kruse, Department of Computer Science, Technical University of Braunschweig, Bültenweg 74 / 75, D-38106 Braunschweig, Germany.

Received June 25, 1991; accepted March 12, 1993.

International Journal of Approximate Reasoning 1993; 9:283–314

655 Avenue of the Americas, New York, NY 10010 0888-613X/93/$6.00

environment from the available symbolic and quantitative approaches to partial ignorance [1] is chosen to be appropriate.

First of all, we must clarify how to interpret the notion of *data* within our approach, and which kinds of imperfectness we intend to consider. On a level of abstraction sufficient for a large number of applications in the field of knowledge-based systems we expect data to characterize the state of an object (*obj*) with respect to underlying relevant frame conditions (*cond*). Regarding the exhaustive formal specification of these frame conditions we assume that it is possible to characterize *obj* by an element state(*obj*, *cond*) of a well-defined set *Dom*(*obj*) of distinguishable object states. *Dom*(*obj*) denotes the domain of the object type attached to *obj*, and is usually called the *universe of discourse* or *frame of discernment* of *obj* with respect to *cond*. Then, considering imperfect knowledge, we are interested in the problem that the original characterization of state (*obj*, *cond*) is not available due to a lack of information about *obj* and *cond*. As an example suppose that *cond* merely permits use of statements like "state(*obj*, *cond*) $\in$ Char(*obj*, *cond*)", where Char(*obj*, *cond*) is a subset of *Dom*(*obj*) and therefore called an *imprecise characterization of obj* w.r.t. *cond*. If, for instance, a fully reliable digital thermometer shows an outside temperature of 25°C, then we may define $\mathrm{Dom}(obj) \overset{Df}{=} [-80, 70]$ as the domain of realistic outside temperatures and just state that the actual outside temperature *obj*, considering the frame conditions *cond* (location of thermometer, date, and time of measurement), is given by $\mathrm{state}(obj, cond) \in [24.5, 25.5]$, i.e., the corresponding characterization is formalized by $\mathrm{Char}(obj, cond) \overset{Df}{=} [24.5, 25.5]$.

For any domain D we call each subset $A \subseteq D$ a *characteristic* of D, in the special case $|A| = 1$ an *elementary* or *precise* characteristic of D. As we postulate the validity of the *closed world assumption* $\mathrm{state}(obj, cond) \in \mathrm{Dom}(obj)$, the application of the *empty* characteristic $\mathrm{Char}(obj, cond) = \emptyset$ is acceptable, but has to be viewed as a contradiction w.r.t *cond*.

Note that $\mathrm{Char}(obj, cond) \subseteq D$ does not contain any information about preferences between the elements of Char(*obj*, *cond*); especially it does not express that they have the same chance to be the unknown original value of state(*obj*, *cond*). Hence the restriction to imprecision as one kind of imperfect knowledge is not satisfactory for real life applications. Whenever the adequate representation of our imperfect knowledge on state (*obj*, *cond*) requires more than a characterization $\mathrm{Char}(obj, cond) \subseteq \mathrm{Dom}(obj)$, namely the incorporation of preferences between the elements of Char(*obj*, *cond*), we interpret these preferences as a hint for the existence of *consideration contexts* that refine the underlying frame conditions *cond*. The corresponding contexts are induced by the presupposition of additional conditions that lead to competing, perhaps (partial) contradicting characterizations of state$(obj, cond)$. Obviously we then deal with a second kind of imperfect knowledge that is different to imprecision. We want to call it *conflict* or *competition*. The combined occurrence of imprecision and conflict in data reflects *vagueness*, and the description of state$(obj, cond)$, based on vague knowledge, a *vague characteristic* of *obj* w.r.t. *cond*. We use vagueness in a narrow sense. Note that notions regarding this topic are not consistent in literature. For a discussion see, for example, [7].

On the formal level a vague characteristic is specified by a function $\gamma: C \to 2^D$, where C denotes a set of consideration contexts and D the underlying domain. The values $\gamma(c)$, $c \in C$, should be viewed as context-dependent (imprecise) characterizations of state$(obj, cond)$.

A simple example for the application of vague characteristics occurs when throwing a marked die. Define $D \overset{Df}{=} \mathrm{Dom}(obj) \overset{Df}{=} \{1, 2, \ldots, 6\}$ to be the appropriate universe of discourse for the characterization of the outcome *obj* of this experiment that is expected to be random. We suppose that there is the positive (subjective) probability p of obtaining an even number when playing the die, and the (subjective) probability $1 - p$ of obtaining an odd number, respectively.

From the context model's point of view we have to distinguish between exactly two contexts, which are c_1 (even number) and c_2 (odd number). Hence, we define the vague characteristic $\gamma: C \to 2^D$, where $C \overset{Df}{=} \{c_1, c_2\}$, $\gamma(c_1) \overset{Df}{=} \{2, 4, 6\}$ and $\gamma(c_2) \overset{Df}{=} \{1, 3, 5\}$.

Related to the given imperfect knowledge on the die, there is neither the chance of refining c_1, c_2 nor the chance of specializing $\gamma(c_1)$ and $\gamma(c_2)$, because no preference between the elements of $\gamma(c_1)$ and $\gamma(c_2)$, respectively, are available. On the other hand, to include all the information on our die, we have to consider *uncertainty* as a third kind of imperfect knowledge.

Uncertainty refers to an existing preference relation of the contexts that supports a decison-maker in coming to a well-founded decision on the unknown value state$(obj, cond)$, representing the (non-predictable) result of our experiment. The relative weighting of the contexts might base on objective measurements or on subjective valuations. Since we do not investigate qualitative, but solely quantitative approaches, the mathematical theory of measurement provides the appropriate formal environment for the representation of uncertainty aspects. Thus, a vague characteristic $\gamma: C \to 2^D$ is called *valuated*, if it is related to a *context measure space* $(C, \mathfrak{S}, P_C)$, where $P_C(\{c\}) \in \mathbb{R}^+$ quantifies the measure (relevance) of context $c \in C$. For practical reasons we restrict ourselves to the treatment of finite measure spaces, i.e., we assume to have $|C| \in \mathbb{N}$ and $\mathfrak{S} \overset{Df}{=} 2^C$. With respect to our example we define $P_C(\{c_1\}) \overset{Df}{=} p$ and $P_C(\{c_2\}) \overset{Df}{=} 1 - p$.

Note that from the mathematical point of view, in case of $P_C(C) = 1$, the mapping γ is a random set, but obviously we use a different interpretation as for example used in [20].

In the following sections we will show in which way the concept of valuated vague characteristics turns out to be useful for the development of an integrating model of vague and uncertain data. On the one hand the resulting context model may be viewed as an autonomous approach to the handling of imperfect knowledge, but on the other hand it delivers a framework for a clarified formal and semantic comparison of concepts given in Bayes theory [2–7], Dempster–Shafer theory [7–11], the transferable belief model [12], possibility theory [13–15], and fuzzy-set theory [16–19].

Due to the many sides of imperfect knowledge it is of course impossible for the context model to remove the user's task of selecting and specifying the decisive frame conditions *cond*, the corresponding universe of discourse Dom(*obj*), and to define a reasonable context measure space with appropriate vague characteristics. (The same work has, for example, to be done in pure probabilistic approaches by the choice of the underlying probability space.) Nevertheless the context model serves to get a unified theory of imperfect data, assists in a strong representation and interpretation of vague and uncertain data by valuated vague characteristics, supports the realization of important operations on valuated vague characteristics (e.g., specialization, generalization, refinement, coarsening, combination, accumulation, conditioning, data revision), and helps in decision-making based on a corresponding inference mechanism.

2. VALUATED VAGUE CHARACTERISTICS

Since the conceptual motivation of the context model has been presented in the previous section, we are now in the position to introduce the formal definition of valuated vague characteristics.

DEFINITION 1 *Let D be a non-empty* universe of discourse (*frame of discernment, domain of a data type*) *and C a nonempty finite* set of contexts.

$\Gamma_C(D) \overset{Df}{=} \{\gamma \mid \gamma: C \to \Gamma(D)\}$ *is defined to be the* set of all vague characteristics of D w.r.t. C.

Ignoring the contexts, $\Gamma(D) \overset{Df}{=} 2^D = \{A \mid A \subseteq D\}$ *designates the* set of all (imprecise) characteristics of D.

Let $\gamma, \gamma' \in \Gamma_C(D)$, $A \in \Gamma(D)$, and $c \in C$.

(a) γ empty, *iff* $\gamma(C) = \{\gamma(c) \mid c \in C\} = \{\emptyset\}$;

(b) γ elementary, *iff* $(\forall c \in C)(|\gamma(c)| = 1)$;

(c) γ context-precise, *iff* $(\forall c \in C)(|\gamma(c)| \leq 1)$;
(d) γ contradictory, *iff* $(\exists c \in C)(\gamma(c) = \emptyset)$;
(e) γ consistent, *iff* $\bigcap_{c \in C} \gamma(c) \neq \emptyset$;
(f) γ c-correct *w.r.t.* A, *iff* $A \subseteq \gamma(c)$;
(g) γ correct *w.r.t.* A, *iff* $(\forall c \in C)(A \subseteq \gamma(c))$;
(h) γ specialization of γ' (γ' *generalization of* γ, γ *more specific than* γ', γ' *correct w.r.t.* γ), *iff* $(\forall c \in C)(\gamma(c) \subseteq \gamma'(c))$;
(i) $ext_C[A] \in \Gamma_C(D)$, *given by* $(\forall c \in C)(ext_C[A](c) = A)$, *is called the extension of* A *w.r.t.* C;
(j) $\gamma(c)$ *is the c-selection (c-projection) of* γ.

Furthermore we introduce the following abbreviations:

$\Gamma_C^0(D) \overset{Df}{=} \{\gamma \in \Gamma_C(D) | \gamma \text{ not empty}\}$;

$\Gamma_C^1(D) \overset{Df}{=} \{\gamma \in \Gamma_C(D) | \gamma \text{ not contradictory}\}$;

$\Gamma_C^2(D) \overset{Df}{=} \{\gamma \in \Gamma_C(D) | \gamma \text{ consistent and not contradictory}\}$;

$C_\gamma \overset{Df}{=} \{c \in C | \gamma(c) \neq \emptyset\}$ (set of non-contradictory contexts w.r.t. γ).

Remark 2
(a) $\Gamma_C^2(D) \subseteq \Gamma_C^1(D) \subseteq \Gamma_C^0(D) \subseteq \Gamma_C(D)$.
(b) Every property that we have defined for vague characteristics w.r.t. D is also referable to the characteristics $A \in \Gamma(D)$ if their extensions $\text{ext}_C[A]$ fulfill this property w.r.t. any context set C. We call, for example, the characteristic $B \in \Gamma(C)$ *correct* w.r.t. the characteristic $A \in \Gamma(C)$, iff $A \subseteq B$.
(c) γ is *valuated* w.r.t. $(C, \Gamma(C), P_C)$, iff $P_C\colon \Gamma(C) \to \mathbb{R}_0^+$ is a finite measure on C which fulfills $(\forall c \in C)(P_C(\{c\}) > 0)$. In this case $(C, \Gamma(C), P_C)$ is called a *context measure space*.

Valuated vague characteristics show a formal analogy to the concept of *random sets* recommended by Matheron [20, 21] and Nguyen [22], though the semantic differences of the two approaches are made clear by the fact that the co-domains of random sets contain set-valuated data to be considered as indivisible objects, whereas the most important application field of a vague characteristic $\gamma \in \Gamma_C(D)$ refers to its interpretation as an imperfect description of the actual state of an object *obj*, given some frame conditions *cond*, where state(*obj*, *cond*) $\in D$. Hence, random sets specify vague concepts or vague properties, but using $\gamma \in \Gamma_C(D)$ as a vague characteristic, the context-dependent selections $\gamma(c)$, $c \in C$, reflect the imprecision involved in the characterization of state(*obj*, *cond*) with respect to context c. The selections $\gamma(c)$ signal that state(*obj*, *cond*) $\in \gamma(c)$ is supposed to be valid, whenever the individual frame conditions of context c are correct for state(*obj*, *cond*).

Apart from these differences the reader who is familiar with random set theory will find some known basic notations.

From a decision-making point of view we intend to valuate the acceptance degree $\mathrm{ACC}_\gamma(A)$, $A \in \Gamma(D)$, that the proposition state(*obj*, *cond*) $\in A$ is true. If we abstract from the given object state to the underlying elementary characteristic $\mathrm{Orig}_\gamma \overset{Df}{=} \{\mathrm{state}(obj, cond)\}$, the so-called *original* of γ, then γ should be viewed as a simplified representation of a set of context-precise characteristics, into which γ may be specialized, each of them with selections that are possible to equal Orig_γ.

The inherent imprecision of γ does not allow to uniquely determine acceptance degrees $\mathrm{ACC}_\gamma(A)$, $A \in \Gamma(D)$, but to calculate upper and lower bounds for them, like done the similar way in random set theory, the theory of upper and lower probabilities, and Dempster–Shafer theory.

DEFINITION 3 *Let* $\gamma \in \Gamma_C^0(D)$.

$$Sel(\gamma) \overset{Df}{=} \{S \mid S\colon C_\gamma \to D \text{ and } (\forall c \in C_\gamma)(S(c) \in \gamma(c))\}$$

denotes the set of all selectors of γ.

In the special case of an empty characteristic we define $Sel(ext_C[\emptyset]) \overset{Df}{=} \emptyset$.

Remark 4 Let $\gamma \in \Gamma_C^0(D)$ and φ_γ: $\mathrm{Sel}(\gamma) \to \Gamma_C(D)$, determined by $(\forall S \in \mathrm{Sel}(\gamma))(\forall c \in C)((\varphi_\gamma(S))(c) = \{S(c)\})$. $\varphi_\gamma(\mathrm{Sel}(\gamma))$ consists of all context-precise characteristics, to which γ may be specialized.

Obviously each selector $S \in \mathrm{Sel}(\gamma)$ is measurable w.r.t. P_C. S induces the finite measure P_S: $\mathscr{B} \to \mathbb{R}_0^+$, $P_S(A) \overset{Df}{=} P_C(S^{-1}(A))$ related to the measurable space $(D, \mathscr{B})$, where $\mathscr{B}$ is expected to be an appropriate σ-field w.r.t. D.

If $\mathrm{Orig}_\gamma \in \Gamma(D)$ is the original of $\varphi_\gamma(S)$, then $P_C(S^{-1}(A))$ quantifies the measure of all contexts c of C, for which there exists a data element $a \in A$, so that $\varphi_\gamma(S)$ is c-correct w.r.t. $\{a\}$. This is the measure of all contexts that do not contradict the proposition $\mathrm{Orig}_\gamma \subseteq A$. For that reason we call $P_C(S^{-1}(A))$ the acceptance degree of A w.r.t. S.

The undermentioned definition generalizes the notion of an acceptance degree from selectors to the underlying vague characteristics.

DEFINITION 5 *Let* $\gamma \in \Gamma_C^0(D)$ *be valuated w.r.t.* $(C, \Gamma(C), P_C)$. *Furthermore let* $\mathscr{B}$ *be an appropriate* σ*-field w.r.t.* D, *and* $A \in \mathscr{B}$.

$$\underline{ACC}_\gamma(A) \overset{Df}{=} \inf\{P_C(S^{-1}(A)) \mid S \in Sel(\gamma)\}$$

is the minimum acceptance degree (necessity),

$$\overline{ACC}_\gamma(A) \overset{Df}{=} \sup\{P_C(S^{-1}(A)) \mid S \in Sel(\gamma)\}$$

is the maximum acceptance degree (possibility)

of A w.r.t. γ.

As shown by the following proposition, it is no problem to support an efficient computation of acceptance degrees:

PROPOSITION 6 *Let* $\gamma \in \Gamma_C^0(D)$ *be valuated w.r.t.* $(C, \Gamma(C), P_C)$. *Then, for all* $A \in \mathscr{B} \setminus \{\emptyset\}$,

$$\underline{ACC}_\gamma(A) = P_C(\{c \in C | \emptyset \neq \gamma(c) \subseteq A\});$$

$$\overline{ACC}_\gamma(A) = P_C(\{c \in C | \emptyset \neq \gamma(c) \cap A\}).$$

Let $\mathrm{Orig}_\gamma \subseteq D$ be an unknown original and $\gamma \in \Gamma_C^0(D)$ its corresponding vague characterization, valuated w.r.t. a context measure space $(C, \Gamma(C), P_C)$. Furthermore suppose to have $\underline{\mathrm{ACC}}_\gamma(A)$ and $\overline{\mathrm{ACC}}_\gamma(A)$, where $a \in \mathscr{B}$ refers to an appropriate measure space $(D, \mathscr{B})$.

Then, we are asked to justify how to come to a decision w.r.t. Orig_γ. Such decision-making aspects will be investigated in section 5, when different types of operations on vague characteristics have been discussed.

3. OPERATIONS ON VAGUE CHARACTERISTICS

Up to now we have considered the representation, interpretation, and valuation of imperfect knowledge by the context model. In this section we introduce the most important operations on vague characteristics, which are specialization (i.e., conditioning, data revision), combination, accumulation, refinement, and coarsening. The mentioned operations are indispensable to provide a model for reasoning in the presence of partial ignorance.

In the first instance we restrict ourselves to the investigation of operations on imprecise characteristics. Suppose $A_i, B_i \in \Gamma(D_i)$, $i = 1, 2, \ldots, n$, and a function f: $\times_{i=1}^n \Gamma(D_i) \to \Gamma(D)$. As the correctness of A_i w.r.t. B_i for every $i = 1, 2, \ldots, n$ does not necessarily imply the correctness of $f(A_1, \ldots, A_n)$ w.r.t. $f(B_1, \ldots, B_n)$, we obviously have to be careful in accepting only reasonable operations on characteristics. Hence we introduce the concepts of correctness- and contradiction-preserving mappings, respectively.

DEFINITION 7 *Let* $D_1, D_2, \ldots, D_n, D$ *be frames of discernment and* f: $\times_{i=1}^n \Gamma(D_i) \to \Gamma(D)$ *a function.*

(a) *f is called* correctness-preserving, *iff* $f(A_1, \ldots, A_n) \subseteq f(B_1, \ldots, B_n)$ *for all* A_i, B_i *with* $A_i \subseteq B_i \subseteq D_i$, $i = 1, 2, \ldots, n$.

(b) *f is called* contradiction-preserving, *iff*

$$(\forall A_1, \ldots, A_n)((\exists i \in \{1, \ldots, n\})(A_i = \emptyset) \Rightarrow f(A_1, \ldots, A_n) = \emptyset)$$

Remark 8

(a) Let $A_i, B_i \in \Gamma(D_i)$, $A_i \subseteq D_i$, $B_i \subseteq D_i$, $i = 1, \dots, n$, and $f\colon \times_{i=1}^{n} \Gamma(D_i) \to \Gamma(D)$. If $(\forall i \in \{1, 2, \dots, n\})(B_i$ correct w.r.t. $A_i)$ and f correctness-preserving, then $f(B_1, \dots, B_n)$ is correct w.r.t. $f(A_1, \dots, A_n)$.

(b) Each function $f\colon \times_{i=1}^{n} D_i \to D$ with $D_1, D_2, \dots, D_n, D \neq \emptyset$ induces a correctness- and contradiction-preserving mapping $\bar{f}\colon \times_{i=1}^{n} \Gamma(D_i) \to \Gamma(D)$, given by $\bar{f}(A_1, \dots, A_n) \overset{Df}{=} f(A_1 \times A_2 \times \cdots \times A_n)$.

(c) The Boolean algebra $(\Gamma(D), \cap, \cup, \mathscr{C})$ consists of the two correctness-preserving operators $\cap$, $\cup$, while $\mathscr{C}$ is not correctness-preserving. Furthermore only $\cap$ is contradiction-preserving.

(d) Compositions of correctness-preserving (contradiction-preserving) mappings are themselves correctness-preserving (contradiction-preserving).

Let us now change over to operations on vague characteristics $\gamma_i \in \Gamma_C(D_i)$, $i = 1, \dots, n$, that refer to a common context measure space $(C, \Gamma(C), P_C)$. We assume that the operations on $\gamma_1, \dots, \gamma_n$ are induced by the *extension* $\underset{\sim}{f}$ of mappings $f\colon \times_{i=1}^{n} \Gamma(D_i) \to \Gamma(D)$ to the given set C of contexts.

DEFINITION 9 *Let* $f\colon \times_{i=1}^{n} \Gamma(D_i) \to \Gamma(D)$ *be a function and* $\gamma_i \in \Gamma_C(D_i)$, $i = 1, 2, \dots, n$. *The vague characteristic* $\underset{\sim}{f}(\gamma_1, \dots \gamma_n) \in \Gamma_C(D)$, *defined by* $(\underset{\sim}{f}(\gamma_1, \dots \gamma_n))(c) \overset{Df}{=} f(\gamma_1(c), \dots \gamma_n(c))$ *is called* the vague image of $(\gamma_1, \dots \gamma_n)$ under f.

In the following subsections we apply the concept of vague images to different types of operations on imperfect data.

3.1. Specialization Concepts: Context Conditioning vs. Data Revision

Let $\gamma \in \Gamma_C(D)$ be the specification of a vague observation w.r.t. an elementary characteristic $\mathrm{Orig}_\gamma \in \Gamma(D)$.

Suppose that further evidence on Orig_γ—to be formalized by a vague characteristic $\nu \in \Gamma_C(D)$—becomes available.

If we claim the correctness of ν w.r.t. Orig_γ, then there are two important justifiable specializations of γ by ν, which are context conditioning and data revision, respectively. Both are structure-preserving, but data-adjusting operations on cooperating vague characteristics referred to the same set of contexts and the same domain.

Context conditioning is mainly related to the competition property within the vague characteristic γ, because it restricts the set C_γ of non-contradictory contexts to exactly those contexts $c \in C$, where the selection $\nu(c)$ is correct w.r.t. all possible specializations of the given imprecise observation $\gamma(c)$ into elementary characteristics of D.

On the other hand data revision is referred to the imprecision property of γ. Data revision does not question the acceptance of contexts on principle, but specializes the selections $\gamma(c)$ based on the assumed correctness of ν w.r.t. Orig_γ.

Hence data revision specializes imprecise observations, whereas context conditioning is limited to the alternatives either to leave imprecise observations unchanged or to make them contradictory.

Definition 10 *Let* $\gamma, \nu \in \Gamma_C(D)$.

(1) $\gamma|\nu: C \to \Gamma(D), (\gamma|\nu)(c) \overset{Df}{=} \begin{cases} \gamma(c), & \textit{iff } \gamma(c) \subseteq \nu(c) \\ \emptyset, & \textit{otherwise} \end{cases}$
is called context conditioning of γ by ν.

(2) $\gamma_\nu: C \to \Gamma(D), \gamma_\nu(c) \overset{Df}{=} \gamma(c) \cap \nu(c)$
is named data revision of γ by ν.

Remark 11

(a) Context conditioning and data revision may be interpreted as extensions of operations on imprecise) characteristics:
Define

$$\mathrm{cond}: \Gamma(D) \times \Gamma(D) \to \Gamma(D), \mathrm{cond}(A, B) \overset{Df}{=} \begin{cases} A, & \text{iff } A \subseteq B \\ \emptyset, & \text{otherwise} \end{cases}$$

$$\mathrm{rev}: \Gamma(D) \times \Gamma(D) \to \Gamma(D), \mathrm{rev}(A, B) \overset{Df}{=} A \cap B.$$

Then for all $\gamma, \nu \in \Gamma_C(D)$ we obtain $\gamma|\nu = \underset{\sim}{\mathrm{cond}}(\gamma, \nu)$ and $\gamma_\nu = \underset{\sim}{\mathrm{rev}}(\gamma, \nu)$.

It should be emphasized that the functions cond and rev are contradiction-preserving, but only rev fulfills the property of correctness-preservation. This fact entails the anomaly of the context conditioning concept that it tends to manipulate the valuation of $\gamma|\nu$ by rejection of contexts $c \in C$ with allocated selection $\gamma(c)$ of minor specificity. For this reason we have to tolerate the restrictive applicability of the context conditioning concept, where in the special case of context-precise characteristics (which are known to be suitable for the representation of discrete probability distributions) context-conditioning coincides with data revision.

(b) Let $\gamma, \nu \in \Gamma_C(D), \emptyset \neq A \subseteq D, c \in C$.
If γ is c-correct w.r.t. A and ν c-correct w.r.t. A, then γ_ν and $\gamma|\nu$ are c-correct w.r.t. A.

3.2. Coordination Concepts: Combination vs. Accumulation

Whenever we specialize a vague characteristic $\gamma \in \Gamma_C(D)$ by another vague characteristic $\nu \in \Gamma_C(D)$ we may view ν as a vague fact that helps us in finding a more specific vague observation of the underlying elementary characteristic $\mathrm{Orig}_\gamma \in \Gamma(D)$.

While specializations are always referred to cooperating vague characteristics with the same domain and the same set of contexts, we now consider operations on competing vague characteristics with a common domain, but disjoint sets of contexts, namely the structure-modifying combination and the structure-preserving accumulation of vague characteristics.

Let $\gamma_i \in \Gamma_{C_i}(D)$, $i = 1, \ldots, n$, be vague characteristics that are related to the same original characteristic, i.e., $\mathrm{Orig}_{\gamma_1} = \mathrm{Orig}_{\gamma_2} = \cdots = \mathrm{Orig}_{\gamma_n}$. A combination of $\gamma_1, \ldots, \gamma_n$ consists of two separable actions, which are the structure-modifying coordination of the corresponding context sets C_1, $C_2, \ldots, C_n$ to a single context set C, and the coordination of all selections of $\gamma_1, \gamma_2, \ldots, \gamma_n$ w.r.t. this integrating context set C by application of an appropriate combination function.

Obviously the most general coordination of contexts is reached when we define C to be the cartesian product of $C_1, \ldots, C_n$.

Hence we state the following:

DEFINITION 12 *Let $\gamma_i \in \Gamma_{C_i}(D)$, $i = 1, 2, \ldots, n$, be vague characteristics of D w.r.t. C_i. Assume that $C_1, C_2, \ldots, C_n$ are pairwise disjoint. Furthermore let $C \overset{Df}{=} C_1 \times C_2 \times \cdots \times C_n$, and let comb: $(\Gamma(D))^n \to \Gamma(D)$ be a correctness- and contradiction-preserving combination function.*
Furthermore we define

$$\overline{comb}: \Gamma(D^n) \to \Gamma(D),$$

$$\overline{comb} \overset{Df}{=} comb(\pi_1(A), \ldots, \pi_n(A)),$$

where $\pi_i(A)$ denotes the projection of A into the i-th coordinate space. The vague characteristic $\underset{\sim}{\overline{comb}}(\gamma_1 \otimes \cdots \otimes \gamma_n) \in \Gamma_C(D^n)$,

$$(\gamma_1 \otimes \cdots \otimes \gamma_n)((c_1, \ldots, c_n)) \overset{Df}{=} \gamma_1(c_1) \times \cdots \times \gamma_n(c_n)$$

is called the combination of $\gamma_1, \ldots, \gamma_n$ w.r.t. comb.

Remark 13

(a) Let $\varnothing \neq A_i \subseteq D_i$, $\gamma_i \in \Gamma_{C_i}(D)$, $c_i \in C_i$, $i = 1, 2, \ldots, n$.
If $(\forall i \in \{1, 2, \ldots, n\})(\gamma_i\ c_i$-correct w.r.t. $A_i)$,
then $\underset{\sim}{\overline{\mathrm{comb}}}(\gamma_1 \otimes \cdots \otimes \gamma_n)$ is $(c_1, \ldots, c_n)$-correct w.r.t. $\widetilde{\mathrm{comb}}\,(A_1, \ldots, A_n)$.

(b) The set-theoretical intersection is an example of a well-known conjunctive combination function, but note that the set-theoretical union does not fulfill the contradiction-preserving property. Nevertheless we may use a modified disjunctive combination function, defined by

$$f_{\sqcup}: \Gamma(D) \times \Gamma(D) \to \Gamma(D),$$
$$f_{\sqcup}(A, B) \overset{Df}{=} \begin{cases} A \cup B, & \text{iff } A \neq \emptyset \text{ and } B \neq \emptyset \\ \emptyset, & \text{iff } A = \emptyset \text{ or } B = \emptyset \end{cases}$$

(c) If $\gamma_i \in \Gamma_{C_i}(D)$, $i = 1, 2, \ldots, n$, are valuated w.r.t. $(C_i, \Gamma(C_i), P_{C_i})$, then $\widetilde{\text{comb}}(\gamma_1 \otimes \cdots \otimes \gamma_n)$ is valuated w.r.t. $(C, \Gamma(C), P_C)$, where $P_C \overset{Df}{=} \otimes_{i=1}^{n} P_{C_i}$, i.e., $P_C(C_1^* \times \cdots \times C_n^*) = \prod_{i=1}^{n} P_{C_i}(C_i^*)$ for all $C_i^* \subseteq C_i$, $i = 1, 2, \ldots, n$.
Note that the application of the product measure P_C requires the typical measure-theoretical independence assumption.

The combination of competing vague characteristics is a structure-modifying operation, because it generates a new set of contexts and therefore—applying a combination function—modified imprecise characteristics as selections within the new contexts. Combination is motivated by structural relationships between the involved context measure spaces.

An alternative approach to the coordination of vague characteristics refers to the introduction of a structure-preserving operation that restricts itself to the union of the given context sets. In this case we suppose irrelative context measure spaces, but apart from that the same frame conditions as mentioned for combination.

DEFINITION 14 *Let $\gamma_i \in \Gamma_{C_i}(D)$, $i = 1, 2, \ldots, n$, be vague characteristics of D w.r.t. C_i. Furthermore let $C_1, C_2, \ldots, C_n$ be pairwise disjoint, and define $C \overset{Df}{=} \bigcup_{i=1}^{n} C_i$.*

The vague characteristic $\oplus_{i=1}^{n} \gamma_i \in \Gamma_C(D)$, given by

$$\bigoplus_{i=1}^{n} \gamma_i \overset{Df}{=} \gamma_1 \oplus \cdots \oplus \gamma_n, \qquad (\gamma_1 \oplus \cdots \oplus \gamma_n)(c) \overset{Df}{=} \gamma_i(c),$$

$$\textit{iff } c \in C_i, \qquad i = 1, \ldots, n,$$

is called the accumulation of $\gamma_1, \ldots, \gamma_n$.

Remark 15 If $\gamma_i \in \Gamma_{C_i}(D)$, $i = 1, 2, \ldots, n$, are valuated w.r.t $(C_i, \Gamma(C_i), P_{C_i})$, then $\oplus_{i=1}^{n} \gamma_i$ is valuated w.r.t. $(C, \Gamma(C), P_C)$, where $P_C \overset{Df}{=} \oplus_{i=1}^{n} P_{C_i}$ and $\oplus_{i=1}^{n} P_{C_i}(C^*) \overset{Df}{=} \sum_{i=1}^{n} P_{C_i}(C^* \cap C_i)$.

3.3. Refinement and Coarsening

As we have already pointed out, the main purpose of the concept of vague characteristics $\gamma \in \Gamma_C(D)$ is to specify vague observations of state-dependent attribute values related to an object under consideration. One main aspect for the specification of vague characteristics is the choice of the domain D and the context set C, which depends on the available information about the given object in its actual state. It is obvious that a more detailed registration of all important frame conditions permits a refinement of C and D, respectively. Refinement does not mean that we add a number of contexts (like done by the connection of contexts) or supply further elementary characteristics (which could be interpreted as a contradiction to our closed work assumption), but rather the disintegration of each context $c \in C$ into subcontexts, and the division of data elements $d \in D$ into finer pieces, respectively. For this reason we formalize the concepts of refinement and coarsening (as the reverse operation) within the context model.

DEFINITION 16 *Let C, C^* be two context sets. C^* is a refinement of C, i.e., $C \sqsubset_\varrho C^*$, iff there exists a surjective mapping ϱ: $C^* \to C$.*

In this case ϱ is called a context-reduction-function. *The measure space $(C^*, \Gamma(C^*), P_{C^*})$ is called the* refinement of $(C, \Gamma(C), P_C)$, *iff C^* is a refinement of C and $(\forall c \in C)(P_C(\{c\}) = P_{C^*}(\{c^* \in C^* | c = \varrho(c^*)\}))$.*

DEFINITION 17 *Let C and C^* be context sets, ϱ: $C^* \to C$ a context-reduction-function and D a domain. Dependent from D and ϱ we define the following mappings:*

(a) Ref_ϱ: $\Gamma_C(D) \to \Gamma_{C^*}(D)$,
$(\forall \gamma \in \Gamma_C(D))(\forall c^* \in C^*)(Ref_\varrho(\gamma)(c^*) = \gamma(\varrho(c^*)))$;

(b) $Coarse_\varrho$: $\Gamma_{C^*}(D) \to \Gamma_C(D)$,
$(\forall \gamma^* \in \Gamma_{C^*}(D))(\forall c \in C)(Coarse_\varrho(\gamma^*)(c) = \bigcup\{\gamma^*(c^*) | c^* \in C^* \wedge c \in \varrho(c^*)\})$.

Let $\gamma \in \Gamma_C(D)$ and $\gamma^* \in \Gamma_{C^*}(D)$.

$Ref_\varrho(\gamma)$ *is* called context-refinement of γ w.r.t. ϱ.

$Coarse_\varrho(\gamma^*)$ *is the* context-coarsening of γ^* w.r.t. ϱ.

PROPOSITION 18 *Let $C \sqsubset_\varrho C^*$, $c \in C$, $c^* \in C^*$, $\gamma \in \Gamma_C(D)$, $\gamma^* \in \Gamma_{C^*}(D)$, and $\varnothing \neq A \subseteq D$.*

(a) *γ c-correct w.r.t. A* $\Leftrightarrow$
$(\forall c^* \in C^*)(\varrho(c^*) = c \Rightarrow$ Ref$_\varrho(\gamma)$ c^*-correct w.r.t. $A)$

(b) *γ c^*-correct w.r.t. A* $\Rightarrow$
$(\forall c \in C)(\varrho(c^*) = c \Rightarrow$ Coarse$_\varrho(\gamma^*)$ c-correct w.r.t. $A)$

DEFINITION 19 *Let D, D^* be two domains. D^* is a refinement of D, i.e., $D \sqsubset_\varrho D^*$, iff there exists a surjective data-reduction-function δ: $D^* \to D$. Related to an arbitrary context set C we define the following mappings:*

(a) Ref_δ: $\Gamma_C(D) \to \Gamma_C(D^*)$,
$(\forall \gamma \in \Gamma_C(D))(\forall c \in C)(Ref_\delta(\gamma)(c) = \delta^{-1}(\gamma(c)))$;

(b) $Coarse_\delta$: $\Gamma_C(D^*) \to \Gamma_C(D)$,
$(\forall \gamma^* \in \Gamma_C(D^*))(\forall c \in C)(Coarse_\delta(\gamma^*)(c) = \delta(\gamma^*(c)))$.

Let $\gamma \in \Gamma_C(D)$ and $\gamma^* \in \Gamma_C(D^*)$.
$Ref_\delta(\gamma)$ is called data-refinement of γ w.r.t. δ.
$Coarse_\delta(\gamma^*)$ is the data-coarsening of γ^* w.r.t. δ.

PROPOSITION 20 *Let $D \sqsubset_\varrho D^*$, $c \in C$, $\gamma \in \Gamma_C(D)$, $\gamma^* \in \Gamma_C(D^*)$, and $\varnothing \neq A \subseteq D$, $\varnothing \neq A^* \subseteq D^*$.*

(a) γ *c-correct w.r.t.* $A \Leftrightarrow \mathrm{Ref}_\delta(\gamma)$ *c-correct w.r.t.* $\delta^{-1}(A)$

(b) γ^* *c^*-correct w.r.t.* $A^* \Rightarrow \mathrm{Coarse}_\delta(\gamma^*)$ *c-correct w.r.t.* $\delta(A^*)$.

Remark 21 The proposed definitions coincide with the corresponding operations suggested by Shafer [9] connected with the application of belief functions. Furthermore it should be pointed out that even nonmonotonic reasoning may be handled within the context model if we, for example, operate on vague characteristics by a composition of context refinement and data revision.

4. MASS DISTRIBUTIONS

In connection with uncertainty modeling considerations we have shown the efficient computation of the minimum/maximum acceptance degree of a characteristic $A \in \Gamma(D)$ to be a generalization of the inaccessible original $\mathrm{Orig}_\gamma \in \Gamma(D)$ of a vague characteristic valuated w.r.t. a context measure space $(C, \Gamma(C), P_C)$:

(a) $\underline{\mathrm{ACC}}_\gamma(A) = P_C(\{c \in C | \varnothing \neq \gamma(c) \subseteq A\})$;

(b) $\overline{\mathrm{ACC}}_\gamma(A) = P_C(\{c \in C | \gamma(c) \cap A \neq \varnothing\})$.

The given representation of $\underline{\mathrm{ACC}}_\gamma(A)$ and $\overline{\mathrm{ACC}}_\gamma(A)$ is directly referred to the underlying contexts. As an alternative—motivated by the additivity property of measures—it is promising to use a representation based on valuations of the characteristics contained in the set $\gamma(C)$ of all selections of γ.

DEFINITION 22 *Let $\gamma \in \Gamma_C(D)$ be valuated w.r.t. $(C, \Gamma(C), P_C)$. m_γ: $\Gamma(D) \to \mathbb{R}_0^+$, $m_\gamma(A) \overset{Df}{=} P_C(\{c \in C | \gamma(c) = A\}$ is called the* mass distribution of γ.

PROPOSITION 23 *Let $\gamma \in \Gamma_C^0(D)$ be valuated w.r.t. $(C, \Gamma(C), P_C)$. For all $A \in \Gamma(D)$ we obtain*

(a) $\underline{\mathrm{ACC}}_\gamma(A) = \sum_{B \in \gamma(C):\ \emptyset \neq B \subseteq A} m_\gamma(B)$,

(b) $\overline{\mathrm{ACC}}_\gamma(A) = \sum_{B \in \gamma(C):\ A \cap B \neq \emptyset} m_\gamma(B)$.

The following propositions show how to get the resulting mass distributions of the most important operations on valuated vague characteristics by means of the mass distributions of the participating operands. Furthermore the essence of well-known concepts of Dempster–Shafer theory with an analogous concept of mass distributions is made evident.

PROPOSITION 24 (mass distribution of specializations) *Let $\gamma, \nu \in \Gamma_C(D)$ be valuated w.r.t. $(C, \Gamma(C), P_C)$. Moreover let $A \in \Gamma(D)$.*

(a) $m_{\gamma|\nu}(A) = \begin{cases} P_C(\{c \in C \mid A = \gamma(c) \subseteq \nu(c)\}), & \text{iff } A \neq \emptyset \\ P_C(\{c \in C \mid \gamma(c) \subsetneq \nu(c)\}), & \text{iff } A = \emptyset \end{cases}$

(*context conditioning*)

(b) $m_{\gamma\nu}(A) = m_{\gamma \mathbin{\underset{\sim}{\cap}} \nu}$ (*data revision*)

Remark 25

(a) In the special case $(\forall c \in C)(\emptyset \neq A \subseteq \nu(c))$ we obtain $m_{\gamma|\nu}(A) = m_\gamma(A)$, but there is no general dependency between $m_{\gamma|\nu}$ and m_γ, m_ν.

(b) The data revision of mass distributions turns out to be compatible with Dempster's rule of conditioning [9], if we disregard the normalization factor due to the application of probability measures in Dempster–Shafer-theory. Note that there is no justification of a normalization within the context model.

PROPOSITION 26 (mass distributions of coordinations) *Let $\gamma_i \in \Gamma_{C_i}^0(D)$, $i = 1, 2, \ldots, n$ be valuated w.r.t. $(C_i, \Gamma(C_i), P_{C_i})$, where $C_1, C_2, \ldots, C_n$ are pairwise disjoint sets of contexts.*

Define $C \overset{Df}{=} C_1 \times C_2 \times \cdots \times C_n$, $C' \overset{Df}{=} C_1 \cup C_2 \cup \cdots \cup C_n$, and choose a combination function comb: $(\Gamma(D))^n \to \Gamma(D)$.
Let $A \subseteq D$.

(a) $m_{\underset{\sim}{comb}(\gamma_1 \otimes \cdots \otimes \gamma_n)}(A) = \sum_{A_1, \ldots, A_n:\ comb(A_1, \ldots, A_n) = A} \prod_{i=1}^n m_{\gamma_i}(A_i)$

(*combination*)

(b) $m_{\gamma_1 \oplus \cdots \oplus \gamma_n}(A) = \sum_{i=1}^n m_{\gamma_i}(A)$ (*accumulation*)

Remark 27 The combination of mass distributions of vague characteristics coincides with Dempster's rule of combination [9] if normalization conditions are ignored.

PROPOSITION 28 (mass distributions of refinements) *Let $\gamma \in \Gamma_C(D)$ be valuated w.r.t. $(C, \Gamma(C), P_C)$. Suppose to have a context-reduction-function*

$$\varrho: C^* \to C$$

and a data-reduction-function δ: $D^* \to D$. *Let* $A \in \Gamma(D)$ *and* $A^* \in \Gamma(D^*)$, *respectively.*

(a) $m_{\mathrm{Ref}_\varrho(\gamma)}(A) = m_\gamma(A)$ (*context refinement*)

(b) $$m_{\mathrm{Ref}_\delta(\gamma)}(A^*) = \begin{cases} m_\gamma(\delta(A^*)), & \textit{iff } A^* \in \{\delta^{-1}(A) \mid A \in \Gamma(D)\} \\ 0, & \textit{otherwise} \end{cases}$$ (*data refinement*)

Remark 29

(a) Let $\nu \in \Gamma_{C^*}(D)$ be valuated w.r.t. the refinement $(C^*, \Gamma(C^*), P_{C^*})$ of $(C, \Gamma(C), P_C)$. Then, for all $A \in \Gamma(D)$, $m_\nu(A) = P_{C^*}(\{c^* \in C^* \mid \nu(c^*) = A\})$. Let $c_1^*, c_2^* \in \{c^* \in C^* \mid \nu(c^*) = A\}$. Since in general $(\mathrm{Coarse}_\varrho(\nu))(\varrho(c_1^*)) \neq (\mathrm{Coarse}_\varrho(\nu))(\varrho(c_2^*))$, with the exception of special cases there is no direct representation of $m_{\mathrm{Coarse}_\varrho(\nu)}$ by m_ν.

(b) Let $\nu \in \Gamma_C(D^*)$ be valuated w.r.t. $(C, \Gamma(C), P_C)$. For all $A \in \Gamma(D)$, the following relationship is valid: $m_{\mathrm{Coarse}_\delta(\nu)}(A) = P_C(\{c \in C \mid \mathrm{Coarse}_\delta(\nu)(c) = A\}) = P_C(\{c \in C \mid \delta(\nu(c)) = A\}) \geq P_C(\{c \in C \mid \nu(c) = \delta^{-1}(A)\}) = m_\nu(\delta^{-1}(A))$, because $\nu(c) \subseteq \delta^{-1}(\delta(\nu(c)))$. For this reason there is in general no direct representation of $m_{\mathrm{Coarse}_\delta(\nu)}$ by m_ν.

5. DECISION-MAKING ASPECTS

The task of decision-making as the final step of an inference process based on the concept of valuated vague characteristics concerns the formalization of someone's betting behavior with respect to the truth of statements of the type $\mathrm{Orig}_\gamma \subseteq A$, $A \in \Gamma(D)$, where Orig_γ is interpreted as the (unknown) original of the given valuated vague characteristic $\gamma \in \Gamma_C^0(D)$ which specifies a vague observation of Orig_γ w.r.t. the domain D and the context space $(C, \Gamma(C), P_C)$.

Minimizing the betting risk, exhausting the whole information about Orig_γ contained in γ and $(C, \Gamma(C), P_C)$, and trying to be as fair as possible, we are caused to provide the following betting behavior:

Let B_0, B_1 be two betters and S_0, S_1 their stakes, where $S_0 + S_1 = P_C(C)$.

B_0 bets on $\mathrm{Orig}_\gamma \subseteq A$ versus $\mathrm{Orig}_\gamma \not\subseteq A$, B_1 vice versa. B_0 accepts the bet, iff $S_0 \leq \underline{\mathrm{ACC}}_\gamma(A)$ and $S_1 \geq \overline{\mathrm{ACC}}_\gamma(D \setminus A)$. On the other hand B_1 accepts the bet, iff $S_1 \leq \underline{\mathrm{ACC}}_\gamma(D \setminus A)$ and $S_0 \geq \overline{\mathrm{ACC}}_\gamma(A)$.

In most situations the carefulness of both betters will prevent us from coming to a decision. Formally spoken, a decision is only possible if γ is a context-precise characteristic.

As an example consider the case $P_C(C) = 1$, where γ induces the discrete probability distribution $\{(d, P_C(\gamma^{-1}(\{d\}))) \mid d \in \bigcup\{\gamma(c) \mid c \in C\}\}$ on D.

The mentioned decision problem motivates the introduction of additional restrictive assumptions that enable us to calculate a single acceptance degree $\mathrm{ACC}_\gamma(A) = \overline{\mathrm{ACC}}_\gamma(A) = \underline{\mathrm{ACC}}_\gamma(A)$ as the foundation for the required decision.

A promising point of attachment is the application of the insufficient reason principle that is a well-known concept of Bayes theory. Transferred to our model it says that we have to assume a uniform distribution on each selection $\gamma(c)$ of the vague characteristic γ under consideration. The justification of this assumption rests on the fact that with respect to the information available by γ and $(C, \Gamma(C), P_C)$ it is not possible to favor one of the elementary characteristics contained in $\gamma(c)$ related to any chosen context $c \in C$. Otherwise we could use a modified specification of the vague observation of Orig_γ.

Obviously the consideration of the phenomenon of imprecision motivates the application of the insufficient reason principle. Conflict, on the other hand, interpreted as the occurrence of competing observation contexts, canonically leads to the generalized insufficient reason principle, which is—for example—conform to the corresponding axiomatically induced principle within the transferable belief model [12].

The following definition states the underlying bet function.

DEFINITION 30 *Let D be a domain with $D \subseteq \mathbb{R}^n$ or $|D| \in \mathbb{N}$. Let $\gamma \in \Gamma_C^0(D)$ be valuated w.r.t. $(C, \Gamma(C), P_C)$, fulfilling $(\forall c \in C)(\gamma(c) \in \mathscr{B})$, where $\mathscr{B}$ is an appropriate σ-field w.r.t. D. Furthermore suppose to have the probability space $(D, \mathscr{B}, P_D)$, where P_D induces the uniform distribution on D.*

We call

$$bet_\gamma : \mathscr{B} \to [0,1],$$

$$bet_\gamma(A) \stackrel{Df}{=} \left[P_C(C_\gamma)\right]^{-1} \cdot \sum_{\substack{c \in C: \\ \gamma(c) \cap A \neq \emptyset}} P_C(\{c\}) \cdot P_D(A | \gamma(c))$$

the bet function of γ w.r.t. $(C, \Gamma(C), P_C)$.

Remark 31 The betting behavior induced by the generalized insufficient reason principle may be described as follows: Let B_0, B_1 be two betters and S_0, S_1 their stakes, where $S_0 + S_1 = 1$. B_0 bets on $\mathrm{Orig}_\gamma \subseteq A$ versus $\mathrm{Orig}_\gamma \not\subseteq A$, B_1 vice versa. B_0 and B_1 accept the bet, iff $S_0 = \mathrm{bet}_\gamma(A)$ and $S_1 = 1 - \mathrm{bet}_\gamma(A)$.

5.1. The "Spoiled Sandwich" Effect

An important requirement for the acceptability of any model of partial ignorance is its conformity with fundamental principles of plausible reasoning that serve as a basis for rational decision-making. One of these well-known principles is the so-called sandwich metaphor [23, 24] that states the following rule:

For any proposition $A \in \Gamma(D)$ and any assumption or evidence $E \in \Gamma(D)$, $E \neq \emptyset$, we expect to calculate

$$\begin{aligned}&\min\{\text{Belief}(A|E), \text{Belief}(A|D \setminus E)\}\\ &\leq \text{Belief}(A)\\ &\leq \max\{\text{Belief}(A|E), \text{Belief}(A|D \setminus E)\}.\end{aligned}$$

The motivation for this principle—translated to decision situations—refers to the idea that nobody needs to engage in unnecessary knowledge acquisition or experimentation [24]:

"If a person would choose the same action for every possible outcome of an experiment, then he or she ought to choose that action without running the experiment."

As natural as such a principle might look at the first glance, there are still controversal discussions on the question, whether the sandwich principle should be regarded as a general principle for plausible reasoning [24–27]. It is easy to prove that conditioning of probability measures coincides with the sandwich principle, whereas data revision of belief functions by Dempster's rule of conditioning contradicts it [24]. In our opinion the main point regarding the relevance of the sandwich principle is to understand that it is referred to uncertain, but precise data, which means that imprecision aspects are not covered by the sandwich principle.

Let us explain this remark by application of the context model, where decision-making depends on bet functions bet_γ induced by valuated vague characteristics $\gamma \in \Gamma_C(D)$ on an underlying context measure space $(C, \Gamma(C), P_C)$. We notice that γ realizes the strict separation of observation contexts (specified by C) and possible states of the object under consideration (specified by D). From the formal point of view it is obvious, but nevertheless worth mentioning that only contexts are valuated, and that each context holds its individual (imprecise) selection $\gamma(c)$ of object states. This unidirectional dependency between contexts and data values has the essential consequence that conditioning as well as data revision operations based on an evidence E, $E \in \Gamma(D)$, are not expected to affect the relative valuations of non-contradictory (i.e., $\gamma(c) \neq \emptyset$) contexts. Thus, accepting the sandwich principle in general, we have no chance to avoid the necessity of unreasonable changes of relative context valuations. Since the assertion of the sandwich principle is related to the data set D, but not directly to the context set C, we have to accept that its relevance is restricted to the case, where all contexts c are identifiable with their attached selections $\gamma(c)$, which reflects the need for the validity of the sandwich principle for all context-precise characteristics.

As this special kind of vague characteristic is exactly the one used in Bayesian-like applications of the context model, it should not be surprising that context-precise characteristics fit the sandwich principle as shown in the following proposition.

PROPOSITION 3.2 *Let $\gamma \in \Gamma_C^0(D)$ be a context-precise vague characteristic which is valuated w.r.t. $(C, \Gamma(C), P_C)$. Then, for all $A \in \Gamma(D)$ and $E \in \Gamma(D)$, $E \neq \emptyset$, we have*

$$\begin{aligned}
&\min\{bet_{\gamma|ext_C(E)}(A),\, bet_{\gamma|ext_C(D \setminus E)}(A)\} \\
&\leq bet_\gamma(A) \\
&\leq \max\{bet_{\gamma|ext_C(E)}(A),\, bet_{\gamma|ext_C(D \setminus E)}(A)\}.
\end{aligned}$$

5.2. The Three Prisoners Problem

The three prisoners problem is one of the most quoted and discussed examples concerning the applicability of Dempster–Shafer theory, especially Dempster's rules of conditioning and combination, respectively. Controversal viewpoints regarding the solution of this problem are still pending. The problem itself is stated as follows:

Let A_1, A_2, and A_3 be three prisoners. One of the prisoners is chosen by the warden to be executed, the others to be saved.

Assume that the choice is given by random draw among the three candidates. Prisoner A_1 asks the guard to name one of the prisoners (different from A_1) who will be saved, arguing that such information would clearly be of little help to him with respect to his potential fate.

We suppose that the guard (who is expected to know the decision of the warden) names A_2, where no further information is available than the reliability of the given answer. Should A_1's opinion about his fate be modified?

First of all, we analyze the three prisoners problem by application of the context model. The analysis turns out to be quite simple caused by the strong semantics of the concepts used within the context model.

Obviously the three prisoners story consists of two parts, which are the choice of the prisoner to be executed, and the statement of the guard that is influenced by the result of that choice.

Related to the first part we distinguish the observation contexts c_i, $i = 1, 2, 3$, specified by the assumption that c_i corresponds to the execution of A_i. The considered object *obj* of interest is an abstract one which is characterized by the name of the prisoner to be executed, i.e., Dom $(obj) \stackrel{Df}{=} \{A_1, A_2, A_3\}$.

As we know that the choice of the prisoner to be executed is representable by the uniform distribution Dom(*obj*), we obtain the following valuated characteristic:

$$\gamma: C \to \Gamma(D), \qquad C \stackrel{Df}{=} \{c_1, c_2, c_3\}, \qquad D \stackrel{Df}{=} \{A_1, A_2, A_3\};$$
$$\gamma(c_i) \stackrel{Df}{=} \{A_i\}, \qquad i = 1,2,3.$$

The underlying context measure space $(C, \Gamma(C), P_C)$ is determined by $P_C(\{c_i\}) \stackrel{Df}{=} \frac{1}{3}$, $i = 1,2,3$.
Hence the a priori valuation of A_1 to be chosen for execution is

$$\underline{\mathrm{ACC}}_\gamma(\{A_1\}) = \overline{\mathrm{ACC}}_\gamma(\{A_1\}) = \tfrac{1}{3}.$$

Further information becomes available by the answer of the guard and the fact that he does not lie. The integration of the guard's statement induces a refinement $(C^*, \Gamma(C^*), P_C^\alpha)$ of the given context space, which is

$C^* \stackrel{Df}{=} \{c_{12}, c_{13}, c_{23}, c_{32}\}$, where

c_{ij} means that prisoner A_i is to be executed and the guards names A_j.

$$P_{C^*}^\alpha(\{c_{12}\}) = \tfrac{1}{3}\alpha, \qquad P_{C^*}^\alpha(\{c_{13}\}) = \tfrac{1}{3}(1-\alpha),$$
$$P_{C^*}^\alpha(\{c_{23}\}) = \tfrac{1}{3}, \qquad P_{C^*}^\alpha(\{c_{32}\}) = \tfrac{1}{3}.$$

The underlying context-reduction-function is defined by

$$\varrho: C^* \to C,$$
$$\varrho(c_{12}) \stackrel{Df}{=} \varrho(c_{13}) \stackrel{Df}{=} c_1,$$
$$\varrho(c_{23}) \stackrel{Df}{=} c_2,$$
$$\varrho(c_{32}) \stackrel{Df}{=} c_3.$$

Note that $P_C(\{c_i\}) = P_{C^*}(\{c^* \in C^* | c_i = \varrho(c^*)\})$.
The value of α reflects that we have no information about the decision behavior of the guard, if A_1 is the prisoner who will be executed.
Since $\alpha \in [0,1]$, we have changed over to the family $\{(C^*, \Gamma(C^*), P_{C^*}^\alpha) | \alpha \in [0,1]\}$ of context measure spaces.

Besides the refinement of the context measure space we observe refinement of the domain, because the answer of the guard offers more additional data.

The refined domain is $D^* \overset{Df}{=} \{A_1, A_2, A_3\}^2$, where $(A_i, A_j) \in D^*$, $i, j \in \{1, 2, 3\}$, means that prisoner A_i will be executed and the guard names A_j. As the corresponding data-reduction function we obtain δ: $D^* \to D$, $\delta((A_i, A_j)) \overset{Df}{=} A_i$, $i, j = 1, 2, 3$.

Now we are in the position to include the evidence E based on the statement of the guard. On the formal level the fact that A_2 will be served is expressed by the data set

$$E \overset{Df}{=} \{A_1, A_2, A_3\} \times \{A_2\}.$$

Hence, we define the following valuated vague characteristic:

$$\begin{aligned}
&\gamma^*: C^* \to \Gamma(D^*),\\
&\gamma^*(c_{12}) \overset{Df}{=} \{(A_1, A_2)\},\\
&\gamma^*(c_{13}) \overset{Df}{=} \{(A_1, A_3)\},\\
&\gamma^*(c_{23}) \overset{Df}{=} \{(A_2, A_3)\},\\
&\gamma^*(c_{32}) \overset{Df}{=} \{(A_3, A_2)\}.
\end{aligned}$$

Data revision by the extension $\varepsilon \overset{Df}{=} \mathrm{ext}_C[E]$ of E w.r.t. C delivers

$$\begin{aligned}
&\gamma_\varepsilon^*: C^* \to \Gamma(D^*)\\
&\gamma_\varepsilon^*(c_{12}) = \{(A_1, A_2)\},\\
&\gamma_\varepsilon^*(c_{13}) = \varnothing,\\
&\gamma_\varepsilon^*(c_{23}) = \varnothing,\\
&\gamma_\varepsilon^*(c_{32}) = \{(A_3, A_2)\}.
\end{aligned}$$

Considering D^* the event that A_1 will be executed is given by $A \overset{Df}{=} \{A_1\} \times \{A_1, A_2, A_3\}$.
Thus $\mathrm{ACC}_{\gamma_\varepsilon^*}(A) = \overline{\mathrm{ACC}}_{\gamma_\varepsilon^*}(A) = \frac{1}{3}\alpha$,
and $\mathrm{bet}_{\gamma_\varepsilon^*}(A) = \frac{1}{3}\alpha(\frac{1}{3} + \frac{1}{3}\alpha)^{-1}$, i.e., $\mathrm{bet}_{\gamma_\varepsilon^*}(A) \in [0, \frac{1}{2}]$,
where the assumption that the guard flips a fair coin when he knows that A_1 will be executed (i.e., $\alpha = \frac{1}{2}$) leads to

$$\mathrm{bet}_{\gamma_\varepsilon^*}(A) = \tfrac{1}{3}.$$

To summarize our results, the context model's solution of the three prisoners problem confirms Pearl's analysis [24, 26] which is obtained through a Bayesian approach. Pearl does not accept the solution by intuitive application of belief functions [26] that delivers bel(A_1 will be executed | guard names A_2) = $\frac{1}{2}$.

It should be emphasized that also the context model does not endorse that solution, but it furthermore clarifies the problems that occur by an incautious application of Dempster's rules of combination and conditioning, respectively.

The context model's analysis does not require any a priori decision of using Bayes theory, Dempster–Shafer theory, or, for example, the transferable belief model. It does nothing more than to strictly transform what we exactly know about the three prisoners problem into a pleasant formal environment. The most relevant step of this transformation is due to the definition of the vague characteristic γ^*: $C^* \to \Gamma(D^*)$ which is valuated w.r.t. context measure spaces $(C^*, \Gamma(C^*), P_{C^*}^{\alpha})$, $\alpha \in [0,1]$. Since γ^* arises to be context-precise, having pairwise disjoint context-dependent selections $\gamma^*(c^*)$, each element of the context space C^* is identified by exactly one element of D^*, which is the typical property of a Bayesian-like analysis within the context model.

On the other hand the general Dempster–Shafer analysis is not restricted to elementary selections $\gamma^*(C^*)$, but also accepts imprecise context-dependent data. Thus, the three prisoners problem is a good example for the inevitability of a Bayesian-like numerical analysis, which means that a serious Dempster-Shafer analysis should therefore come to the same result. Indeed, using the mass distribution m_{γ^*} induced by γ^*, which is m_{γ^*}: $\Gamma(D^*) \to [0,1]$, $m_{\gamma^*}(A) \overset{Df}{=} P_{C^*}^{\alpha}(\{c^* \in C^* | \gamma(c^*) = A\})$, $\alpha \in [0,1]$, we calculate the correct solution.

It should not be concealed that the Dempster–Shafer-like analysis of the three prisoners problem by the context model questions (like Pearl [24]), whether the straightforward application of belief functions can serve as a basis for decision-making, because mass distributions and belief functions (which are similar to normalized minimum acceptance degrees) are not fundamental, but induced concepts holding imprecision to be removed by bet functions.

6. WHAT IS THE DEMPSTER–SHAFER–SMETS MODEL?

In this section we outline essential remarks regarding the interpretation of the Dempster–Shafer–Smets model (DSSM) within the context model. The Dempster–Shafer–Smets model reflects Smets' "non-probabilistic" view of Dempster–Shafer theory, modified by the idea of transferring basic belief assignments due to data revision [12]. For this reason Smets himself calls it the transferable belief model.

First of all we summarize the basic ideas of the DSSM as they are introduced by Smets in [12].

Let Ω be a nonempty finite set, called the frame of discernment, and $(2^{\Omega}, \cap, \cup, \mathscr{C})$ the induced Boolean algebra of propositions on Ω. The description of our subjective personal judgment that propositions of 2^{Ω} are true, is called a credal state on the propositional space $(\Omega, 2^{\Omega})$. It is formalized by a mapping val: $\Omega \to \{\text{true}, \text{false}\}$, where $|\text{val}^{-1}(\{\text{true}\})| = 1$, which means that the truth lies in exactly one of the elements of Ω.

Suppose that $\{\omega_0\} = \text{val}^{-1}(\{\text{true}\})$. A proposition $A \in 2^{\Omega}$ is called true, iff $\omega_0 \in A$.

A basic principle of the DSSM postulates that the impact of an evidence on ω_0 consists of allocating parts of an initial unitary amount of belief among the propositions of $\Gamma(\Omega)$. The mentioned allocation is defined by a basic belief assignment m: $2^{\Omega} \to [0, 1]$, fulfilling the condition $\sum_{A:\, A \in 2^{\Omega}} m(A) = 1$.

If further evidence becomes available and implies that the truth is in a subset B of Ω, i.e., $\omega_0 \in B$, then the belief $m(A)$ initially allocated to $A \in \Gamma(\Omega)$ is transferred to $A \cap B$.

The resulting basic belief assignment is defined by Dempster's rule of conditioning:

$$m_B(A)\colon 2^{\Omega} \to [0, 1],$$

$$m_B(A) \overset{Df}{=} \begin{cases} c \cdot \sum\limits_{X:\, X \subseteq \Omega \setminus B} m(A \cup X), & \text{iff } A \subseteq B \\ 0, & \text{iff } A \subsetneq B, \end{cases}$$

where $c \overset{Df}{=} 1$ (Smets' proposal; normalization dropped),

or $c \overset{Df}{=} \left(1 - \sum\limits_{X:\, X \subseteq \Omega \setminus B} m(X)\right)^{-1}$ (Shafer's proposal; normalization).

After this short introduction we change over to the context model: The mentioned credal state is an epistemic construct and relative to our knowledge. It is based on a not necessarily unknown finite set C of competing contexts of consideration that support the specification of imprecise observations of the not directly accessible elementary characteristic $\{\omega_0\} \subseteq \Omega$, where ω_0 is a state-dependent attribute value of the object under consideration.

In the DSSM the valuated vague observation of $\{\omega_0\}$ is specified by a basic belief assignment m: $2^{\Omega} \to [0, 1]$. Within the context model the mentioned observation of $\{\omega_0\}$ is formalized by a valuated vague characteristic $\gamma \in \Gamma_C(\Omega)$ with respect to the chosen set C of contexts and a valuating measure $P_C: C \to \mathbb{R}_0^+$. Hence, the induced mass distribution m_γ: $2^\Omega \to \mathbb{R}_0^+$, $m_\gamma(A) \overset{Df}{=} P_C(\{c \in C | \gamma(c) = A\})$, $A \subseteq \Omega$, corresponds to the basic belief assignment m of the DSSM.

If we know that the truth is in a subset B of Ω, i.e., $\omega \in B$, we postulate the correctness of B w.r.t. $\mathrm{Orig}_\gamma = \{\omega_0\}$.

So we have a data revision by the extension of B w.r.t. C. The revised vague characteristic is $\gamma_{\mathrm{ext}_C[B]} \in \Gamma_C(\Omega)$, defined by $\gamma_{\mathrm{ext}_C[B]}(c) \overset{Df}{=} \gamma(c) \cap (\mathrm{ext}_C[B])(c) = \gamma(c) \cap B$.

Thus, for all $A \in 2^\Omega$, we calculate

$$\begin{aligned} m_{\gamma_{\mathrm{ext}_C[B]}}(A) &= P_C\left(\left\{c \in C | \gamma_{\mathrm{ext}_C[B]}(c) = A\right\}\right) \\ &= P_C(\{c \in C | \gamma(c) \cap B = A\}) \\ &= \sum_{X:\, X \cap B = A} P_C(\{c \in C | \gamma(c) = X\}) \\ &= \sum_{X:\, X \cap B = A} m_\gamma(X) = \sum_{X:\, X \subseteq \Omega \setminus B} m_\gamma(A \cup X), \end{aligned}$$

which coincides with Smet's proposal of Dempster's unnormalized rule of conditioning.

7. EXAMPLE: SOLUTION OF THE UNRELIABLE ALARM PARADIGM

One of the examples that Smets [12] introduced to illustrate situations where the DSSM leads to results different from those of its contenders—especially the Bayesian model—is referred to the valuation of propositions on the good behavior of a system in an environment of partial ignorance. The object to be characterized by a tuple of state-dependent attribute values consists of the system itself, a sensor, and an alarm bell. For this reason we consider the attribute values SYSTEM, SENSOR, and ALARM with their domains DOM(SYSTEM) $\overset{Df}{=}$ {on, down}, DOM(SENSOR) $\overset{Df}{=}$ {working, broken} and DOM(ALARM) $\overset{Df}{=}$ {ringing, quiet}, respectively.

The following rules give a description of the general dependencies between the three attributes:

R_1: Probability(SENSOR = working) = 0.8;
Probability(SENSOR = broken) = 0.2;

R_2: If SENSOR = working
then if SYSTEM = on
then ALARM = quiet
else ALARM = ringing
else ALARM $\in$ {quiet, ringing};

R_3: SENSOR and SYSTEM are independent from each other.

No further information is available.

Suppose that the alarm bell is ringing at a given time. What is our degree of belief that the system is down at this time?

We will investigate this problem from two different points of view, both of them applications of the context model: one approach is based on the DSSM, the other one is adopted from Bayes theory.

7.1. Solution Based on the Dempster–Shafer–Smets Model

Obviously our frame of discernment is defined by

$$\Omega \overset{Df}{=} \mathrm{DOM(SYSTEM)} \times \mathrm{DOM(SENSOR)} \times \mathrm{DOM(ALARM)}.$$

With respect to the given rules R_1, R_2, and R_3, we establish that only R_1 supports valuation aspects. It permits the distinction between two contexts of consideration, which are the context w of a working sensor and the context b of a broken sensor. Hence we define the context space $(C, \Gamma(C), P_C)$, where $C \overset{Df}{=} \{w, b\}$ and the valuation measure P_C: $\Gamma(C) \to \mathbb{R}_0^+$, $P_C(\{w\}) \overset{Df}{=} 0.8$, $P_C(\{b\}) \overset{Df}{=} 0.2$ (motivated by R_1) are given.

For reasons of abbreviation we introduce the following characteristics on Ω:

$$
\begin{aligned}
&\mathrm{ON} && \overset{Df}{=} \{\mathrm{on}\} \times \mathrm{DOM(SENSOR)} \times \mathrm{DOM(ALARM)},\\
&\mathrm{DOWN} && \overset{Df}{=} \Omega \setminus \mathrm{ON}\\
&\mathrm{WORKING} && \overset{Df}{=} \mathrm{DOM(SYSTEM)} \times \{\mathrm{working}\} \times \mathrm{DOM(ALARM)},\\
&\mathrm{BROKEN} && \overset{Df}{=} \Omega \setminus \mathrm{WORKING}\\
&\mathrm{RINGING} && \overset{Df}{=} \mathrm{DOM(SYSTEM)} \times \mathrm{DOM(SENSOR)} \times \{\mathrm{ringing}\},\\
&\mathrm{QUIET} && \overset{Df}{=} \Omega \setminus \mathrm{RINGING}.
\end{aligned}
$$

Suppose $\omega_t \in \Omega$ to be the unknown original state of our object at time t. From this we specify our vague observation of $\{\omega_t\}$ by a vague characteristic $\gamma \in \Gamma_C(\Omega)$. A priori we know that $\gamma(w) =$ WORKING and $\gamma(b) =$ BROKEN.

R_2 induces the specification of a context-independent and time-independent observation of ω_t, that is $R \overset{Df}{=} \Omega \setminus \{(\text{on}, \text{working}, \text{ringing}), (\text{down}, \text{working}, \text{quiet})\}$.

Further context-independent evidence is available at time t, since we have the additional information that the alarm bell is ringing. This evidence is formalized by the characteristic $E \overset{Df}{=}$ RINGING.

R and E are assumed to be correct for Orig_γ, i.e. $\{\omega_t\} = \text{Orig}_\gamma \subseteq R \cap E$. The fuzzifications of R and E w.r.t. C, designated by $\text{ext}_C(R)$ and $\text{ext}_C(E)$, respectively, induce a data revision of γ.

Using $\gamma_{\text{rev}} \overset{Df}{=} (\gamma_{\text{ext}_C[R]})_{\text{ext}_C[E]}$ we obtain:

$$\begin{aligned}
\gamma_{\text{rev}}(w) &= \gamma(w) \cap (\text{ext}_C[R])(w) \cap (\text{ext}_C[E])(w) \\
&= \{(\text{down}, \text{working}, \text{ringing})\}, \\
\gamma_{\text{rev}}(b) &= \{(\text{on}, \text{broken}, \text{ringing}), (\text{down}, \text{broken}, \text{ringing})\};
\end{aligned}$$

$$m_{\gamma_{\text{rev}}}: \Gamma(\Omega) \to \mathbb{R}_0^+,$$

$$m_{\gamma_{\text{rev}}}(A) = \begin{cases} 0.8, & \text{iff } A = \gamma_{\text{rev}}(w) \\ 0.2, & \text{iff } A = \gamma_{\text{rev}}(b) \\ 0, & \text{otherwise}; \end{cases}$$

$$\underline{\text{ACC}}_{\gamma_{\text{rev}}}(\text{DOWN}) = 0.8; \qquad \underline{\text{ACC}}_{\gamma_{\text{rev}}}(\text{ON}) = 0;$$

$$\overline{\text{ACC}}_{\gamma_{\text{rev}}}(\text{DOWN}) = 1; \qquad \overline{\text{ACC}}_{\gamma_{\text{rev}}}(\text{ON}) = 0.2;$$

$$\begin{aligned}
\text{bet}_{\gamma_{\text{rev}}}(\text{DOWN}) &= \sum_{\substack{c \in C: \\ \gamma_{\text{rev}}(c) \cap \text{DOWN} \neq \varnothing}} P_C(\{c\}) \cdot \frac{|\text{DOWN} \cap \gamma_{\text{rev}}(c)|}{|\gamma_{\text{rev}}(c)|} \\
&= 0.9; \\
\text{bet}_{\gamma_{\text{rev}}}(\text{ON}) &= 0.1.
\end{aligned}$$

Obviously there is a high degree of belief (0.9) that the system is down if the alarm bell is ringing.

It has to be emphasized that the DSSM approach is based on a restricted set of contexts. So we have to decide whether the neglection of inaccessible, but probably important valuation dependencies is acceptable or not.

In our example it seems to be questionable to content ourselves with the two chosen contexts of consideration (sensor working/broken), as we expect an influence of the system's reliability (on/down) to a justified decision on the system's state at time t.

7.2. Solution by Application of Bayes Theory

The final remark of the previous section indicates that the vague characteristic $\gamma \in \Gamma_C(\Omega)$ may be viewed as a context coarsening of another vague characteristic $\gamma^* \in \Gamma_{C^*}(\Omega)$, arranged by a context reduction mapping $\varrho: C^* \to C$.

In fact the starting point of a pure probabilistic (Bayesian) approach to the unreliable alarm paradigm is the definition of a refined context set C^*, where each element corresponds to one possible world of consideration, that is an elementary characteristic of our frame of discernment.

Therefore we define $C^* \overset{Df}{=} \Omega$, $\gamma^*: C^* \to \Gamma(\Omega)$, given by $(\forall c^* \in C^*)$ $(\gamma^*(c^*) \overset{Df}{=} \{c^*\})$, and $\varrho: C^* \to C$,

$$\varrho(c^*) \overset{Df}{=} \begin{cases} w, & \text{iff } c^* \in \text{WORKING} \\ b, & \text{iff } c^* \in \text{BROKEN.} \end{cases}$$

Obviously we have $\text{Coarse}_\varrho(\gamma^*) \equiv \gamma$, which means that the DSSM-related vague characteristic γ is a context coarsening of the Bayesian-related vague characteristic γ^*.

It credits the Bayesian approach to choose a context set that coincides with all possible dependencies between the attribute values of the considered object, even if the available information on the actual state of that object does not support a unique valuation of the mentioned dependencies. As one consequence there is no need for context refinements if new valuation information occurs.

On the other hand, the Bayesian approach has to handle the problem that incomplete valuation information lacks in finding a unique measure $P_{C^*}: \Gamma(C^*) \to \mathbb{R}_0^+$.

Referred to the unreliable alarm paradigm we get the following valuation information:

R_1: $P_{C^*}(\text{WORKING}) = 0.8$; $P_{C^*}(\text{BROKEN}) = 0.2$;

R_2: $P_{C^*}(\text{QUIET}|\text{WORKING} \cap \text{ON}) = 1$,
$P_{C^*}(\text{RINGING}|\text{WORKING} \cap \text{DOWN}) = 1$;

R_3: $P_{C^*}(\text{DOWN} \cap \text{WORKING}) = P_{C^*}(\text{DOWN}) \cdot P_{C^*}(\text{WORKING})$.

Let $\pi_1 \overset{Df}{=} P_{C^*}(\text{DOWN})$ and $\pi_2 \overset{Df}{=} P_{C^*}(\text{RINGING}|\text{BROKEN})$. Then we obtain the parametric probability distribution $P[\pi_1, \pi_2] \overset{Df}{=}$

$\{(c^*, P_{C^*}[\pi_1, \pi_2](\{c^*\}) | c^* \in C^*\}$, where $P_{C^*}[\pi_1, \pi_2]$ is defined by the undermentioned tabular:

	ON		DOWN	
	RINGING	QUIET	RINGING	QUIET
WORKING	0	$0.8(1-\pi_1)$	$0.8\pi_1$	0
BROKEN	$0.2(1-\pi_1)\pi_2$	$0.2(1-\pi_1)(1-\pi_2)$	$0.2\pi_1\pi_2$	$0.2\pi_1(1-\pi_2)$

The probabilities π_1 and π_2 are not available, i.e., all we know is $\pi_1, \pi_2 \in [0, 1]$. As careful probabilists we consider the family $\mathscr{P} \overset{Df}{=} \{P[\pi_1, \pi_2] | \pi_1, \pi_2 \in [0, 1]\}$ of probability distributions.
Since

$$P_{C^*}(\text{DOWN}|\text{RINGING}) = \begin{cases} \pi_1 \cdot \dfrac{0.8 + 0.2\pi_2}{0.8\pi_1 + 0.2\pi_2}, & \text{iff } \pi_1 \cdot \pi_2 > 0 \\ 0, & \text{otherwise}, \end{cases}$$

we get

$$\inf_{\pi_1, \pi_2 \in [0,1]} \{P_{C^*}(\text{DOWN}|\text{RINGING})\} = 0,$$

$$\sup_{\pi_1, \pi_2 \in [0,1]} \{P_{C^*}(\text{DOWN}|\text{RINGING})\} = 1.$$

In spite of the information that the alarm bell is ringing, we remain in total ignorance about the system's state at time t. So we have no choice but adding further restrictive assumptions to come to a decision within the Bayesian approach.

The typical aid to solve this problem is the application of the insufficient reason principle.

Let $\gamma_{\text{rev}} \overset{Df}{=} (\gamma_{\text{ext}_C[R]})_{\text{ext}_C[E]}$. According to the symbolics used in the previous section we find the following results:

$$\gamma^*_{\text{rev}}(c^*) = \begin{cases} \{c^*\}, & \text{iff } c^* \in R \cap E \\ \varnothing, & \text{otherwise} \end{cases}$$

$$m_{\gamma^*_{\text{rev}}}: \Gamma(\Omega) \to \mathbb{R}_0^+,$$

$$m_{\gamma^*_{\text{rev}}}(A) = \begin{cases} 0.2(1-\pi_1)\pi_2, & \text{iff } A = \{(\text{on}, \text{broken}, \text{ringing})\} \\ 0.2\pi_1\pi_2, & \text{iff } A = \{(\text{down}, \text{broken}, \text{ringing})\} \\ 0.8\pi_1, & \text{iff } A = \{(\text{down}, \text{working}, \text{ringing})\} \\ 0.8(1-\pi_1) & \\ \quad +0.2(1-\pi_2), & \text{iff } A = \varnothing \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} \underline{\mathrm{ACC}}_{\gamma^*_{\text{rev}}}(\text{DOWN}) &= P_{C^*}(\{c^* \in C^* | \varnothing \neq \gamma^*_{\text{rev}}(c^*) \subseteq \text{DOWN}\}) \\ &= 0.2\pi_1\pi_2 + 0.8\pi_1 = \pi_1(0.8 + 0.2\pi_2) \\ \overline{\mathrm{ACC}}_{\gamma^*_{\text{rev}}}(\text{DOWN}) &= P_{C^*}(\{c^* \in C^* | \gamma^*_{\text{rev}}(c^*) \cap \text{DOWN} \neq \varnothing\}) \\ &= \pi_1(0.8 + 0.2\pi_2) \\ \underline{\mathrm{ACC}}_{\gamma^*_{\text{rev}}}(\text{ON}) &= \overline{\mathrm{ACC}}_{\gamma^*_{\text{rev}}}(\text{ON}) = 0.2(1-\pi_1)\pi_2 \\ \text{bet}_{\gamma^*_{\text{rev}}}(\text{DOWN}) &= \pi_1(0.8 + 0.2\pi_2)/\big(1 - m_{\gamma^*_{\text{rev}}}(\varnothing)\big) \\ \text{bet}_{\gamma^*_{\text{rev}}}(\text{ON}) &= 0.2(1-\pi_1)\pi_2/\big(1 - m_{\gamma^*_{\text{rev}}}(\varnothing)\big) \end{aligned}$$

Note that in the context model there is no justification for the normalization of valuations, but the conformity of results with the traditional Bayesian approach is illustrated by

$$P_{C^*}(\text{DOWN}|\text{RING}) = \text{bet}_{\gamma^*_{\text{rev}}}(\text{DOWN}).$$

By application of the insufficient reason principle ($\pi_2 \stackrel{Df}{=} 0.5$) we obtain:

$$\begin{aligned} \text{bet}_{\gamma^*_{\text{rev}}}(\text{DOWN}) &= 0.9\pi_1/\big(1 - m_{\gamma^*_{\text{rev}}}(\varnothing)\big), \\ \text{bet}_{\gamma^*_{\text{rev}}}(\text{ON}) &= 0.1(1-\pi_1)/\big(1 - m_{\gamma^*_{\text{rev}}}(\varnothing)\big). \end{aligned}$$

In opposite to the DSSM the Bayesian approach shows that there is a strict dependency of $\text{bet}_{\gamma^*_{\text{rev}}}(\text{DOWN})$ and $\text{bet}_{\gamma^*_{\text{rev}}}(\text{ON})$ from the general reliability of the system (quantified by π_1).

If we have, for example, a system of high reliability ($\pi_1 \stackrel{Df}{=} 0.01$), then $\text{bet}_{\gamma^*_{\text{rev}}}(\text{DOWN}) = 0.0833$, $\text{bet}_{\gamma^*_{\text{rev}}}(\text{ON}) = 0.9167$, and the decision "*system is on at time t*" has to be preferred. On the other hand, if $\pi_1 \stackrel{Df}{=} 0.99$, then $\text{bet}_{\gamma^*_{\text{rev}}}(\text{DOWN}) = 0.9989$, $\text{bet}_{\gamma^*_{\text{rev}}}(\text{ON}) = 0.0011$, and "*system is down at time t*" is the favorite decision. So the Bayesian approach gives us a hint that we need a quantification of the system's reliability if we want to get a well founded decision.

7.3. Short Comparison: Dempster–Shafer–Smets Model vs. Bayes Theory

The main advantage of the DSSM becomes apparent in its information-related selection of contexts; there is no refinement of contexts, whenever it is not necessary because of a lack of valuating information.

On the other hand, the simplification of consideration by context coarsening has to be justified dependent from the given decision problem, which turns out to be a difficult task in many cases. Since several defenders of the DSSM unfortunately tend to omit the mentioned justification, they often calculate misleading results that factually hide existing decision-making problems. A careful selection of contexts could avoid much of the criticism on the Dempster–Shafer–Smets model.

The Bayesian approach enforces the full structuring of possible dependencies between all attributes of the object under consideration. Therefore it supports a better foundation of the decision-making process, since there is no hidden coarsening of contexts. As a consequence all additional assumptions that extend the disposable information—like, for example, the insufficient reason principle—have to be explicitly defined. The disadvantage of the Bayesian approach lies in the necessity of exhaustively structured contexts, even if only incomplete valuation information is given.

Within the context model the formal differences of the two considered models are the following:

The Bayesian approach deals with a family $\mathscr{P} = \{(C, \Gamma(C), P_C[t]) | t \in T\}$ of context spaces with a fixed set C of contexts and parameterized probability measures $P_C[t]$, $t \in T$.

Additional valuating information induces a restriction of T and therefore the transition to a subset of $\mathscr{P}$.

The DSSM operates on a single context space $(C, \Gamma(C), P_C)$, which has to be refined in the light of new valuating information. Hence the underlying set of contexts is not expected to be fixed.

Nevertheless it should be pointed out that there are also strong connections of the two approaches within the context model: Both approaches make use of vague characteristics to specify vague observations of elementary characteristics. Furthermore they apply the same selected types of operations on vague characteristics (e.g., specialization (conditioning, data revision), combination, refinement, coarsening) and after all the same decision making process by the bet-function-concept of the context model.

8. CONCLUDING REMARKS

In this paper we have introduced the concept of valuated vague characteristics and the context model as an integrating model of vagueness and uncertainty. The motivation for the context model arises from the intention to develop a common formal environment that supports a better understanding and comparison of existing models of partial ignorance to reduce the rivalry between well-known approaches.

Related to a given decision problem the typical structuring used in Bayes theory has turned out to be conform to context-precise valuated characteristics due to sets of exhaustively separated contexts, whereas the typical Dempster–Shafer–Smets structuring tends to coincide with the application of vague characteristics with respect to simplified consideration contexts. This coarsening seems to be helpful for many decision problems of practical interest, but on the other hand it is of course quite dangerous, whenever the choice of consideration contexts is doubtful.

The relationships between the Context Model and Possibility Theory as developed by Dubois and Prade [14] is established by operations on so-called possibility functions induced by information compression mappings $\Pi_{C,D}$: $\Gamma_C(D) \to \mathrm{Poss}(D), \Pi_{C,D}(\gamma) \overset{Df}{=} \pi[\gamma]$, where $\mathrm{Poss}(D) \overset{Df}{=} \{\pi | \pi: D \to \mathbb{R}_0^+ \wedge |\pi(D)| \in \mathbb{N}\}$ and $\pi[\gamma]$: $D \to \mathbb{R}_0^+$; $\pi[\gamma](d) \overset{Df}{=} P_C(\{c \in C | d \in \gamma(c)\})$, for all vague characteristics $\gamma \in \Gamma_C(D)$ that are valuated w.r.t. $(C, \Gamma(C), P_C)$.

Possibility functions have formal relationships to possibility distributions, fuzzy sets, plausibility functions for singletons, and one-point coverages of random sets, but there are differences because of the semantics of the underlying valuated vague characteristics.

The consideration of correctness- and contradiction-preserving mappings, incorporating the additional concept of sufficiency-preservation, offers interesting results regarding the context model's justification of Zadeh's extension principle [29], the principle of minimum specificity [30], and the relevance of the Gödel relation for approximate reasoning with respect to conjunctive systems of inference rules [31]. The corresponding new approach to possibility theory and fuzzy set theory by the context model is topic of separate papers [32, 33].

References

1. Kruse, R., and Siegel, P. (Eds.), *Symbolic and Quantitative Approaches to Uncertainty*, *Lecture Notes in Computer Science* 548, Springer, Berlin, 1991.
2. Savage, L. J., *The Foundations of Statistics*, Wiley, New York, 1954.
3. Finetti, B. de, *Theory of Probability*, *Vol. 1 / 2*, Wiley, New York, 1974–1975.
4. Fishburn, P. C., The axioms of subjective probability, *Stat. Science* 1, 335–358, 1986.
5. Kyburg, H. E. Jr., Bayesian and Non-Bayesian Evidential Updating, *AI* 31, 271–294, 1987.

6. Shafer, G., and Pearl, J., *Readings in Uncertain Reasoning*, Morgan Kaufmann, San Mateo, CA, 1990.

7. Kruse, R., Schwecke, E., Heinsohn, J., *Uncertainty and Vagueness in Knowledge Based Systems—Numerical Methods, Series Artificial Intelligence*, Springer, Berlin, 1991.

8. Dempster, A. P., Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.* 38, 325–339, 1967.

9. Shafer, G., *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976.

10. Shafer, G., Perspectives on the theory and practice of belief functions, *Int. J. Approx. Reasoning* 4, 323–362, 1991.

11. Shafer, G., Rejoinder to "Perspectives on the theory and practice of belief functions," *Int. J. Approx. Reasoning* 6, 445–480, 1992.

12. Smets, P., The transferable belief model: Comparison with Bayesian models, Technical Report, *IRIDIA*, Université Libre de Bruxelles, Belgium, 1989.

13. Zadeh, L. A., Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1, 3–28, 1978.

14. Dubois, D., Prade, H., *Possibility Theory—An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, 1987.

15. Dubois, D., Prade, H., Fuzzy sets in approximate reasoning: Inference with possibility distributions, *Fuzzy Sets and Systems* 40, 143–202, 1991.

16. Zadeh, L. A., Fuzzy sets, *Information and Control* 8, 338–353, 1965.

17. Klir, G. J., Folger, T. A., *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, New York, 1988.

18. Dubois, D., and Prade, H., Fuzzy sets, probability, and measurement, *Eur. J. Operational Res.* 40, 135–154, 1989.

19. Kruse, R., and Meyer, K. D., *Statistics with Vague Data*, Reidel, Dordrecht, 1987.

20. Matheron, G., *Random Sets and Integral Geometry*, Wiley, New York, 1975.

21. Matheron, G., *Estimating and Choosing*, Springer, Berlin, 1989.

22. Nguyen, H. T., On random sets and belief functions, *J. Math. Analysis and Appl.* 65, 531–542, 1978.

23. Aleliunas, R. A., A new normative theory of probabilistic logic, *Proc. 7th Biennial Conf. of the Canadian Society for Computational Studies of Intelligence*, Edmonton, Alta, 67–74.

24. Pearl, J., Reasoning with belief functions: An analysis of compatibility, *Int. J. Approx. Reasoning* 5, 363–390, 1990.

25. Smets, P., Resolving misunderstandings about belief functions, *Int. J. Approx. Reasoning* 6, 321–344, 1992.

26. Pearl, J., Rejoinder to "Reasoning with belief functions: An analysis of compatibility," *Int. J. Approx. Reasoning* 6, 425–443, 1992.

27. Ruspini, E. H., Lowrance, J. D., and Strat, T. M., Understanding evidential reasoning, *Int. J. Approx. Reasoning* 6, 401–424, 1992.

28. Cheeseman, P., In defense of probability, *Proc. 9th IJCAI*, Los Angeles, 1002–1009, 1985.

29. Zadeh, L. A., The concept of a linguistic variable and its application to approximate reasoning, Part I: *Inf. Sciences* 8, 199–249, Part II: *Inf. Sciences* 8, 301–357, Part III: *Inf. Sciences* 9, 43–80, 1975.

30. Dubois, D., and Prade, H., The principle of minimum specificity as a basis for evidential reasoning, in *Uncertainty in Knowledge Based Systems* (B. Bouchon, R. R. Yager, Eds.), Springer, New York, 75–84, 1986.

31. Dubois, D., and Prade, H., Possibility theory as a basis for preference propagation in automated reasoning, *Proc. FUZZIEEE'92*, San Diego, 821–832, 1992.

32. Gebhardt, J., and Kruse, R., A possibilistic interpretation of fuzzy sets by the context model, *Proc. FUZZIEEE'92*, San Diego, 1089–1096, 1992.

33. Gebhardt, J., and Kruse, R., A new approach to possibility theory and fuzzy sets (in preparation), 1992.