On an Information Compression View of Possibility Theory

Jörg Gebhardt and Rudolf Kruse
Department of Computer Science
University of Braunschweig
Bültenweg 74–75
D-38106 Braunschweig, Germany
e-mail: gebhardt@ibr.cs.tu-bs.de

Abstract—**In this paper we consider possibility distributions as information–compressed representations of imperfect characterizations of object states. With the corresponding formalized semantic background, we investigate how to operate on possibility distributions and how to introduce appropriate uncertainty measures for this framework. It turns out that the well–known concept of a possibility measure and an additional certainty measure, which is not its dual necessity measure, are justified as reasonable choices.**

I. Introduction

One basic motivation for possibility distributions is to consider them as imprecise and uncertain descriptions of object states ω_0 of interest. Choosing an appropriate universe of discourse Ω, and stating the closed world assumption $\omega_0 \in \Omega$, a possibility distribution $\pi : \Omega \rightarrow [0,1]$ for ω_0 may be interpreted in the way that, for each $\omega \in \Omega$, $\pi(\omega)$ quantifies the possibility of truth of the proposition "$\omega = \omega_0$".

In the last years there have been many publications on the clarification of the semantics of possibility degrees. Numerical approaches to be mentioned are based, for example, on the epistemic interpretation of fuzzy sets [Zadeh, 1978], the axiomatic view of a theory of possibility [Dubois and Prade, 1988], contour functions of consonant belief functions [Shafer, 1976], one–point–coverages of random sets [Nguyen, 1978], falling shadows in set–valued statistics [Wang, 1983], Spohn's theory of epistemic states [Spohn, 1990], and possibility theory based on likelihoods [Dubois et al. 1993].

In this paper we prefer to relate the semantic background of possibility theory to imperfect characterizations $\Gamma = (\gamma, P_C)$ of ω_0, defined as a multivaluated mapping $\gamma : C \rightarrow 2^{\Omega}$ w.r.t. an underlying (normalized) measure space $(C, \mathfrak{A}, P_C)$.

In section II we introduce this view, which, to some extent, refers to similar considerations that were suggested, for example, in [Strassen, 1964], [Dempster, 1968], and [Kampé de Fériet, 1982], but it is in fact a more general formal and semantic framework for handling imperfect data in knowledge–based systems [Gebhardt and Kruse, 1993a, Gebhardt and Kruse, 1993b]. We then investigate *possibility functions* (non–normalized possibility distributions) as information–compressed representations π_Γ of imperfect characterizations Γ, and briefly consider how to operate on them in this setting. Among other things this leads to an alternative justification of the extension principle [Zadeh, 1975].

In section III we focus our interest on the definition of uncertainty measures that coincide with our view of possibility functions.

As a result we obtain a new approach to the well–known notion of a *possibility measure* [Zadeh, 1978] as an upper uncertainty measure for possibility theory, whereas a specific *certainty measure* – which is *not* its dual necessity measure – seems to be the adequate choice for the corresponding lower uncertainty measure.

II. Information Compression View of Possibility Functions

In this section we introduce possibility functions as information–compressed representations of imperfect characterizations of an object state $\omega_0 \in \Omega$ under consideration, or, in a more general sense, of a non–empty set $\Omega_0 \subseteq \Omega$ of possible object states, where $\Omega_0 = \{\omega_0\}$ occurs as a special case. The starting point of our investigation is a multivalued mapping $\gamma : C \rightarrow 2^{\Omega}$ w.r.t. an underlying normalized measure space $(C, \mathfrak{A}, P_C)$. The basic idea is to view C as a set of competing consideration contexts, each of them chosen to be possibly adequate for an impre-

This work has been partially funded by CEC–ESPRIT III Basic Research Project 6156 (DRUMS II)

0-7803-1896-X/94 $4.00 ©1994 IEEE

cise characterization of Ω_0. Any context $c \in C$, to be specified as a proposition in a common algebra of propositions, delivers its individual imprecise characterization $\gamma(c)$ of Ω_0. At first glance, C might be interpreted as a sample space, which means that exactly one context in C is assumed to be the right choice for characterizing Ω_0, but we do not know which one. If $c \in C$ turns out to be the true context, then $\gamma(c)$ is expected to be *correct* ($\Omega_0 \subseteq \gamma(c)$) and of *maximum specificity* w.r.t. Ω_0 (no proper subset of $\gamma(c)$ is guaranteed to satisfy the correctness condition). Restricting ourselves in the following to the important case of finite context sets, $P_C(\{c\})$ then quantifies the *probability of truth* of context c. On the other hand, dropping the sample space assumption implies that there is no longer the condition of having one and only one true context in C for the characterization of Ω_0. A typical example of this situation is given when C is a set of sensors which imprecisely observe Ω_0. In this case, the interpretation of P_C has to be weakened in the sense that $P_C(\{c\})$ simply reflects an additive *correctness weight*, to be carefully chosen based on specificity and reliability assumptions for context c. An imperfect characterization $\Gamma = (\gamma, P_C)$, $\gamma : C \to 2^\Omega$, is therefore called *α–correct* w.r.t. Ω_0, iff $P_C(\{c \in C | \Omega_0 \subseteq \gamma(c)\}) \geq \alpha$, where $\alpha \geq 0$.

It is obvious that a well–founded and reasonable application of this approach requires to clarify the semantics of correctness weights, how to obtain them in practice, how to consistently operate on imperfect characterizations, and how to justify decision making procedures w.r.t. α–correctness assumptions. But this is out of the scope of this paper. For some basic issues on the mentioned topics we refer to [Gebhardt and Kruse, 1993a], while a more detailed consideration regarding the special problem of combining evidence in various numerical settings is presented in [Gebhardt and Kruse, 1993b]. Using sample spaces of contexts, handling of imperfect characterizations supports, for example, a comparison and alternative justification of concepts that are known from probabilistic reasoning and Dempster–Shafer theory. In this paper we drop the sample space assumption and rather focus on possibility theory, where possibility functions, at least from a pure data–oriented and knowledge–representation point of view (i.e.: ignore the exact semantics of the underlying contexts), coincide with one–point–coverages of random sets [Nguyen, 1978].

Definition 2.1
Let $\Gamma = (\gamma, P_C)$, $\gamma : C \to 2^\Omega$, be an imperfect characterization. Then,

$$\pi_\Gamma : \Omega \to [0,1],$$
$$\pi_\Gamma(\omega) \overset{Df}{=} P_C(\{c \in C \mid \omega \in \gamma(c)\})$$

is called the *induced possibility function of* Γ. Furthermore, let POSS(Ω) denote the set of all induced possibility functions w.r.t. Ω. For $\pi \in$ POSS(Ω), Repr(π) $\overset{Df}{=} \{(\alpha, [\pi]_\alpha) \mid \alpha \in [0,1]\}$ with the *α–cuts* $[\pi]_\alpha \overset{Df}{=} \{\omega \in \Omega \mid \pi(\omega) \geq \alpha\}$, $\alpha \in (0,1]$, $[\pi]_0 \overset{Df}{=} \Omega$, is the *identifying set representation* of π.
Finally, if $\pi_1, \pi_2 \in$ POSS(Ω) and $\pi_1 \leq \pi_2$, then π_1 is called *as least as specific* as π_2.

Presupposing the α–correctness of Γ w.r.t. Ω_0, note that $[\pi_\Gamma]_\alpha$ is the most specific imprecise characterization of Ω_0. For this reason possibility functions are the appropriate choice for representing our imperfect knowledge Γ on Ω_0, whenever α–correctness assumptions are made. In our setting, operating on possibility functions is then characterized as follows: Given an operation $\phi : \bigtimes_{i=1}^{n} 2^{\Omega_i} \to 2^\Omega$ and n possibility functions $\pi_i \in$ POSS(Ω_i), where $\pi_i \equiv \pi_{\Gamma_i}$, $\Gamma_i = (\gamma_i, P_{C_i})$, $\gamma_i : C_i \to 2^{\Omega_i}$, $i = 1, \ldots, n$, Γ_i being imperfect characterizations of (inaccessible) non-empty $\Omega_0^{(i)} \subseteq \Omega_i$, and assuming α–correctness of Γ_i w.r.t. $\Omega_0^{(i)}$, calculate the most specific possibility function $\pi \in$ POSS(Ω) such that $[\pi]_\alpha$ is correct w.r.t. $\phi(\Omega_0^{(1)}, \ldots, \Omega_0^{(n)})$. In fact, defining $\phi[\pi_1, \ldots, \pi_n] \in$ POSS(Ω) by $[\phi[\pi_1, \ldots, \pi_n]]_\alpha \overset{Df}{=} \phi([\pi_1]_\alpha, \ldots, [\pi_n]_\alpha)$, $\alpha > 0$, it turns out that $\pi \equiv \phi[\pi_1, \ldots, \pi_n]$, iff ϕ satisfies a property called *sufficiency–preservation* [Gebhardt and Kruse, 1992].

Definition 2.2
Let $\phi : \bigtimes_{i=1}^{n} 2^{\Omega_i} \to 2^\Omega$ be an operation on imprecise characterizations. ϕ is called *sufficiency–preserving*, iff

(a) $(\exists i \in \{1, \ldots, n\})(A_i = \emptyset) \implies \phi(A_1, \ldots, A_n) = \emptyset$

(b) $\phi(A_1 \cup B_1, \ldots, A_n \cup B_n) = \bigcup \left\{ \phi(C_1, \ldots, C_n) \mid C_j \in \{A_j, B_j\}, j = 1, \ldots, n \right\}$

for all $A_i, B_i \subseteq \Omega_i$, i=1,...,n.

The identity $\pi \equiv \phi[\pi_1, \ldots, \pi_n]$ reflects a generalized form of the *extension principle* [Zadeh, 1975], provable through the semantic background of possibility functions and the application of sufficiency–preserving operations (which are, for instance, intersection, projection, and cylindrical extension, just to mention a few of those that are relevant in the field of possibilistic reasoning systems). In a similar way other basic concepts of possibility theory (e.g.: principle of minimum specificity, projection–combination mechanism [Dubois and Prade, 1991]) are not just principles, but also provable as theorems in our framework. More detailed results regarding the field of possibilistic reasoning are contained in [Gebhardt and Kruse, 1993c].

III. Possibility Measures and Certainty Measures

In the previous section we introduced an information–compression view of possibility functions and discussed how to operate on them with respect to the chosen formal and semantic background. The aim of this section is to briefly consider aspects of possibilistic decision making. This means, given a possibility function $\pi \in \mathrm{POSS}(\Omega)$ as an imperfect characterization of a non–empty set $\Omega_0 \subseteq \Omega$ of possible object states, we are searching for a decision D such that $\Omega_0 \subseteq D$ is accepted to be true.

Again, we start with the assumption that there are n (inaccessible) imperfect characterizations $\Gamma_i = (\gamma_i, P_{C_i})$, $\gamma_i : C_i \to 2^{\Omega_i}$ of underlying unknown, non–empty sets $\Omega_0^{(i)} \subseteq \Omega_i$, $i = 1, \ldots, n$, and induced, observable possibility functions π_{Γ_i} such that Γ_i is α–correct w.r.t. $\Omega_0^{(i)}$. Furthermore let $\phi : \bigtimes_{i=1}^{n} 2^{\Omega_i} \to 2^{\Omega}$ be an operation on imprecise characterizations. For any chosen $A \subseteq \Omega$ we then want to calculate

(1) the *minimum* of all correctness degrees α such that the correctness of A w.r.t. $\phi(\Omega_0^{(1)}, \ldots, \Omega_0^{(n)})$ is *certain*,

(2) the *maximum* of all correctness degrees α such that the correctness of A w.r.t. $\phi(\Omega_0^{(1)}, \ldots, \Omega_0^{(n)})$ remains at least *possible*.

Based on these correctness presuppositions, (1) induces the most pessimistic, (2) the most optimistic decision w.r.t. $\phi(\Omega_0^{(1)}, \ldots, \Omega_0^{(n)})$. The following definition formalizes what we intend to achieve:

Definition 3.1
Let $\Gamma_i = (\gamma_i, P_{C_i})$, $\gamma_i : C_i \to 2^{\Omega_i}$, be imperfect characterizations, $\phi : \bigtimes_{i=1}^{n} 2^{\Omega_i} \to 2^{\Omega}$ an operation on imprecise characterizations, and $A \subseteq \Omega$.

(a)

$$\mathrm{Cert}[\phi; \Gamma_1, \ldots, \Gamma_n](A) \stackrel{Df}{=} \inf\Big\{\alpha \mid \Big(\forall (A_1, \ldots, A_n) \in \bigtimes_{i=1}^{n} 2^{\Omega_i}\Big) \Big(\big(\forall i \in \{1, \ldots, n\}\big) (\Gamma_i\ \alpha\text{–correct w.r.t. } A_i) \Longrightarrow (A \text{ correct w.r.t. } \phi(A_1, \ldots, A_n))\Big)\Big\}$$

is called the *certainty degree of A w.r.t. ϕ and* $(\Gamma_1, \ldots, \Gamma_n)$.

(b)

$$\mathrm{Poss}[\phi; \Gamma_1, \ldots, \Gamma_n](A) \stackrel{Df}{=} \sup\Big\{\alpha \mid \Big(\exists (A_1, \ldots, A_n) \in \bigtimes_{i=1}^{n} 2^{\Omega_i}\Big) \Big(\big(\forall i \in \{1, \ldots, n\}\big) (\Gamma_i\ \alpha\text{–correct w.r.t. } A_i) \wedge (A \text{ correct w.r.t. } \phi(A_1, \ldots, A_n))\Big)\Big\}$$

is defined to be the *possibility degree of A w.r.t. ϕ and* $(\Gamma_1, \ldots, \Gamma_n)$.

The following theorem shows that certainty and possibility degrees can directly be calculated with the aid of the induced possibility functions $\pi_{\Gamma_1}, \ldots, \pi_{\Gamma_n}$, without the need of referring to the underlying imperfect characterizations $\Gamma_1, \ldots, \Gamma_n$.

In advance we state a helpful definition of two uncertainty measures, one of which is the well–known possibility measure in possibility theory [Dubois and Prade, 1988], whereas the other one induces a different concept (certainty measure) and therefore *not* occurs as the dual necessity measure.

Definition 3.2
Let $\pi \in \mathrm{POSS}(\Omega)$.

(a) $\mathrm{Cert}[\pi] : 2^{\Omega} \to [0, 1]$,
$\mathrm{Cert}[\pi](A) \stackrel{Df}{=} \inf\{\alpha \mid \emptyset \neq [\pi]_\alpha \subseteq A\}$,
$\inf \emptyset \stackrel{Df}{=} \infty$,

is called the *certainty measure induced by* π.

(b) $\mathrm{Poss}[\pi] : 2^{\Omega} \to [0, 1]$,
$\mathrm{Poss}[\pi](A) \stackrel{Df}{=} \sup\{\alpha \mid [\pi]_\alpha \cap A \neq \emptyset\}$

is the *possibility measure induced by* π.

If $\mathrm{Cert}[\pi](A) = \infty$, then no correctness assumption can ensure the correctness of A w.r.t. the non–empty set Ω_0, imperfectly characterized by π.

Theorem 3.3
Let $\Gamma_i = (\gamma_i, P_{C_i})$, $\gamma_i : C_i \to 2^{\Omega_i}$, be imperfect characterizations, $\phi : \bigtimes_{i=1}^{n} 2^{\Omega_i} \to 2^{\Omega}$ a *sufficiency-preserving* operation, and $A \subseteq \Omega$.

(a) $\text{Cert}[\phi; \Gamma_1, \ldots, \Gamma_n](A) = \text{Cert}[\phi[\pi_{\Gamma_1}, \ldots, \pi_{\Gamma_n}]](A)$

(b) $\text{Poss}[\phi; \Gamma_1, \ldots, \Gamma_n](A) \leq \text{Poss}[\phi[\pi_{\Gamma_1}, \ldots, \pi_{\Gamma_n}]](A)$

(c) If there exists a mapping $\psi : \bigtimes_{i=1}^{n} \Omega_i \to \Omega$ such that $\phi(A_i, \ldots, A_n) = \psi(A_i \times \cdots \times A_n)$, then (b) changes to equality.

IV. Concluding Remarks

In spite of the proposed rich semantic background of possibility functions it has turned out that operations on them as well as decision making based on the concepts of α-correctness, certainty and possibility measures, respectively, can be realized by only considering the induced possibility functions rather than their underlying imperfect characterizations, as far as the sufficiency-preservation property of the involved operators is satisfied. The whole investigation therefore supports a well-founded and effective handling of possibility functions.
For this reason, in cooperation with German Aerospace, the presented approach to possibility theory has successfully been applied in order to develop a software tool (POSSINFER) for possibilistic reasoning in multidimensional spaces of hypotheses [Kruse et al., 1994].

References

[Dempster, 1968] A.P. Dempster, Upper and Lower Probabilities Induced by a Random Closed Interval, Ann. Math. Stat., 39:957-966, 1968.

[Dubois and Prade, 1988] D. Dubois and H. Prade, Possibility Theory, Plenum Press, New York, 1988.

[Dubois and Prade, 1991] D. Dubois and H. Prade, Basic Issues on Fuzzy Sets in Approximate Reasoning, Part 1: Inference with Possibility Distributions, Fuzzy Sets and Systems, 40:143-202, 1991.

[Dubois et al. 1993] D. Dubois, S. Moral, H. Prade, A Semantics for Possibility Theory Based on Likelihoods, CEC-ESPRIT III BRA 6156 DRUMS II, Annual Report, 1993.

[Gebhardt and Kruse, 1992] J. Gebhardt and R. Kruse, A Possibilistic Interpretation of Fuzzy Sets by the Context Model, In Proc. FUZZ-IEEE'92, 1089–1096, San Diego, 1992.

[Gebhardt and Kruse, 1993a] J. Gebhardt and R. Kruse, The Context Model – An Integrating View of Vagueness and Uncertainty, Int. J. of Approximate Reasoning, 9:283–314, 1993.

[Gebhardt and Kruse, 1993b] J. Gebhardt and R. Kruse, A Comparative Discussion of Combination Rules in Numerical Settings, CEC-ESPRIT III BRA 6156 DRUMS II, Annual Report, 1993.

[Gebhardt and Kruse, 1993c] J. Gebhardt and R. Kruse, A New Approach to Semantic Aspects of Possibilistic Reasoning, In M. Clarke, R. Kruse, and S. Moral, editors, Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Lecture Notes in Computer Science, 747, 151–160, Springer, Berlin, 1993.

[Kampé de Fériet, 1982] J. Kampé de Fériet, Interpretation of Membership Functions of Fuzzy Sets in Terms of Plausibility and Belief, In M.M. Gupta, E. Sanchez, editors, Fuzzy Information and Decision Processes, North-Holland, 13–98, 1982.

[Kruse et al., 1994] R. Kruse, J. Gebhardt, and F. Klawonn, Fuzzy Systems, Theory and Application, Wiley, Chichester, 1994. (Translation of the book "Fuzzy-Systeme", Series: Leitfäden und Monographien der Informatik, Teubner, Stuttgart, 1993).

[Nguyen, 1978] H.T. Nguyen, On Random Sets and Belief Functions, J. of Mathematical Analysis and Applications, 65:531–542, 1978.

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

[Spohn, 1990] W. Spohn, A General Non-Probabilistic Theory of Inductive Reasoning, In R.D. Shachter, T.S. Levitt, L.N. Kanal, and J.F Lemmer, editors, Uncertainty in Artificial Intelligence, 149–158, North Holland, Amsterdam, 1990.

[Strassen, 1964] V. Strassen, Meßfehler und Information. Zeitschrift Wahrscheinlichkeitstheorie und verwandte Gebiete, 2:273–305, 1964.

[Wang, 1983] P.Z. Wang, From the Fuzzy Statistics to the Falling Random Subsets, In P.P. Wang, editor, Advances in Fuzzy Sets, Possibility and Applications, 81–96, Plenum Press, New York, 1983.

[Zadeh, 1975] L.A. Zadeh, The Concepts of a Linguistic Variable and its Application to Approximate Reasoning, Information Sciences, 8:199–249, 301–357, 9:43–80, 1975.

[Zadeh, 1978] L.A. Zadeh, Fuzzy Sets as a Basis for a Theory of Possibility, Fuzzy Sets and Systems, 1:3–28, 1978.