Chapter 5
On Possibilistic and Probabilistic Information Fusion

Ronald R. Yager
Iona College, USA

ABSTRACT

This article discusses the basic features of information provided in terms of possibilistic uncertainty. It points out the entailment principle, a tool that allows one to infer less specific from a given piece of information. The problem of fusing multiple pieces of possibilistic information is and the basic features of probabilistic information are described. The authors detail a procedure for transforming information between possibilistic and probabilistic representations, and using this to form the basis for a technique for fusing multiple pieces of uncertain information, some of which is possibilistic and some probabilistic. A procedure is provided for addressing the problems that arise when the information to be fused has some conflicts.

INTRODUCTION

Information used in decision making generally comes from multiple sources. We are interested in the problem of multi-source information fusion in the case when the information provided has some uncertainty. Two important types of sources of information are electro-mechanical sensors and human observers/experts; this is particularly the case in security environments. We note that sensor-provided information generally has a probabilistic type of uncertainty. Human-observer-provided information, which is generally linguistic in nature, typically introduces a possibilistic type of uncertainty. Here we are faced with a problem in which we must fuse information with different modes of uncertainty. Here we shall discuss an approach to attaining capability based on the use of probability-possibility transformation. We first discuss the basic features of information provided in terms of possibilistic uncertainty. We point out the entailment principle, a tool that allows one to infer less specific from a given piece of information. We also discuss the problem of fusing multiple pieces of possibilistic information. We describe a procedure for transforming information between possibilistic and probabilistic representations. We use this to form the basis for a technique for fusing multiple pieces of uncertain information some of which is possibilistic and some probabilistic. We

DOI: 10.4018/978-1-4666-1870-1.ch005
provide a normalization procedure for addressing the problems that arise when the information to be fused has some conflict.

The most well-known uncertainty representation is the probabilistic model. Assume \( V \) is some variable taking its value in the space \( X \). In probabilistic uncertainty we associate with each \( x_i \in X \) a value \( p_i \in [0, 1] \) indicating the probability that \( x_i \) is the value of \( V \). We denote the collection of these as the probability distribution \( P \). For any subset \( A \) of \( X \) the probability that \( V \) lies in \( A \) is

\[
\Pr(A) = \sum_{i: x_i \in A} p_i.
\]

In the probabilistic framework the normalization property that \( \Pr(X) = 1 \) requires that \( \sum_i p_i = 1 \). Probabilistic uncertainty is essentially based on ratio scale information. That is, if we know the ratio \( \frac{p_j}{p_i} \) then we can completely determine the probability distribution. Within the framework of probability theory the Shannon entropy is used to quantify the amount of uncertainty associated with a probability distribution \( P \)

\[
H(p) = -\sum_{i=1}^{n} p_i \log(p_i).
\]

**Possibilistic Uncertainty**

Assume \( V \) is some variable taking its value in the space \( X \). Formally under a possibilistic uncertainty model we associate with each \( x_i \in X \) a value \( \pi_i \in [0, 1] \) indicating the possibility that \( x_i \) is the value of \( V \). We denote the collection of these as the possibilistic distribution \( \Pi \). For any subset \( A \) of \( X \) the possibility that \( V \) lies in \( A \) is

\[
\text{Poss}(A) = \max_{x_i \in A} \pi_i.
\]

In the possibilistic framework the normalization property that \( \text{Poss}(X) = 1 \) requires that \( \max_{i} \pi_i = 1 \). Possibilistic uncertainty is essentially based on ratio scale information. That is, if we know the ratio \( \frac{\pi_j}{\pi_i} \) then we can completely determine the probability distribution. Within the framework of probability theory the Shannon entropy is used to quantify the amount of uncertainty associated with a probability distribution \( P \)

\[
H(p) = -\sum_{i=1}^{n} p_i \log(p_i).
\]

Possibilistic uncertainty often arises from linguistically expressed information via a fuzzy set. Again assume \( V \) is some variable whose domain is the set \( X \). A proposition or datum associated with this variable is a statement of the form \( V \) is \text{Value}. Here \text{Value} is some constraint on the possible values of the variable (Zadeh, 2005). For example if \( V \) is John’s age then \text{Value} could be “young” or “old”, etc. Using the framework of computing with words introduced by Zadeh (1996, 1999) we can express the meaning of \text{Value} in terms of a fuzzy subset \( A \) of \( X \). In the case of John’s age we would define “young” as a fuzzy over ages. Using the fuzzy subset \( A \) we can induce a possibility distribution \( \Pi \) to express our knowledge of the value \( V \) by assigning \( \pi_i = A(x_i) \). Thus the possibility that \( V = x_i \) is the membership grade of \( x_i \) in \( A \).

In passing we want to make one comment regarding the assignment statement \( V \) is \text{Value}. In point of fact what we have denoted as the variable \( V \) can be more precisely viewed as consisting of an attribute and an object. Thus a variable is of the form \text{Attribute} (\text{Object}). For example in the statement John is young, the more precise statement is age of John is young. Here we have a datum triple, attribute – object – value. Nevertheless in the following we shall continue to use the term variable.

Returning to our possibility distribution \( \Pi \) on \( V \). We indicated that for any subset \( A \) we have \( \text{Poss}[A] = \max_{x_i \in A} \pi_i \). We recall that for any subset \( A \) we can associate a function \( A: X \rightarrow [0, 1] \) such that \( A(x_i) = 1 \) if \( x_i \in A \) and \( A(x_i) = 0 \) if \( x_i \notin A \). Using this we can express \( \text{Poss}[A] = \max_{x_i \in A} [A(x_i) \wedge \pi_i] \). This formulation allows us to calculate the possibility of fuzzy sets when \( A(x_i) \in [0, 1] \) instead of \{0, 1\}.

Within the framework of possibility theory the concept of specificity (Yager, 1998; Klir, 2006) plays a role analogous to the role of entropy in probability theory, it measures the uncertainty associated with a possibility distribution. Let \( \text{ind}(A) \)