Adaptive Behavior of Perceptual Information Processing

Atsushi INOUE*
E-mail: ainoue@ececs.uc.edu
Department of Electrical and Computer Engineering and Computer Science,
University of Cincinnati,
Cincinnati, OH 45221-0030 U.S.A.

Anca L. RALESCU †
E-mail: anca@islab.brain.riken.go.jp
Information Synthesis Laboratory, Brain Science Institute, RIKEN,
2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan

Abstract

A robust method of intelligent information processing with respect to user's perception of data, named Perceptual Information Processing (PIP), is being studied. This method uses fuzzy sets as the representation of user perceptions and a soft computing method called Mass Assignment Theory (MAT) as the computational core to manage consistencies between histograms of actual data and the fuzzy sets. This paper describes PIP's adaptive behavior such that user perception is reinforced as it is applied to recognize a data set.

Introduction

A novel information processing method with respect to the user perception of data, named Perceptual Information Processing (PIP), has been proposed in (Inoue and Ralescu 98) and is currently under further study, both at theoretical and application levels. This is analogous to a simple and robust feature extraction method often used for image and signal processing scheme such that certain types of features extracted from an actual data set are used to generate a recognition result.

In PIP fuzzy sets (Zadeh 65) are used to capture and represent a user perception of data. Their use is justified by their dual numeric-symbolic nature, as well as by the fact that subjectivity and bias in the data can be effectively represented and exploited. More than one fuzzy sets can be used to represent a given a collection of data. Therefore one of the issues that arises is that of obtaining the best, with respect to some criterion, of these fuzzy sets. In connection with this issue of central importance is the consistency between the actual data and the fuzzy set used to represent it. In most of the relevant literature (e.g. practically any paper on fuzzy modeling or fuzzy control), this consistency issue is not considered, and fuzzy sets, are adapted in a more or less *ad-hoc* manner.

The main characteristic of PIP is the use of a soft computing (Zadeh 96) method based on Mass Assignment Theory (MAT) (Baldwin 95). This offers the theoretical framework to manage consistency between data, feature histograms, and fuzzy sets used to model perception by an user of the data (Ralescu 97) capturing in this process an individual user's idiosyncrasies. In this respect PIP is similar to artificial neural networks (ANN). However, the level of handling semantics is different such that where ANN establishes a mapping between inputs and outputs by simulating biological neurons PIP also manages consistencies in terms of probability distributions and fuzzy sets and it sets up the stage for deriving hidden relationships between the data.

The focus of this paper is on the adaptive behavior of PIP, similar to the behavior exhibited by ANN. The current perception of data by an user is reinforced in order to get a better recognition result. A small and simple example is taken to illustrate the adaptive behavior of PIP, and an experiment with realistic data sets, i.e. a text corpus, is explained. This experiment is in progress as a part of the dissertation research of the first author.

Mass Assignment Theory (MAT)

Description

Mass assignment theory (MAT) provides a framework for managing consistency between fuzzy sets and probability distributions. For a fixed probability distribution (PD), there are many fuzzy sets (FS) which can correspond to it (1PD-mFS). Conversely, there are many probability distributions corresponding to a particular fuzzy set (1FS-mPD). Both of the consistency management become possi-

Research partially supported by E-volv Adaptive Technology, 10549 Reading Road, Cincinnati, OH 45241-2524 U.S.A.

On leave from ECECS, University of Cincinnati, Cincinnati, OH 45221-0030 U.S.A.

ble due to the formalization of selection rules defined on nested focal elements.

The above one-to-many relations reflect the following:

1. **1D-mFS:** There are many interpretations for a particular data set with certain range of possible interpretation, (e.g. a price of an item can be interpreted differently by different individuals);
2. **1FS-mPD:** Given a particular interpretation of a data set, several actual data sets may fit this interpretation.

Definitions

The basic concepts of MAT are defined as follows:

Definition 1 *Let* $\mathbf{S}$ *be a sample space. A mass assignment (MA)* $\mathbf{m_s}$ *associated to* $\mathbf{S}$ *is a function from the power set* $\mathcal{P}(S)$ *to an interval of real numbers such that*

$$m_s : \mathcal{P}(S) \longrightarrow [0, 1] \quad (1)$$

and

$$\sum_{A \subseteq S} m_s(A) = 1 \quad (2)$$

Definition 2 *A subset* $\mathbf{A}$ *of a sample space* $\mathbf{S}$ *is called a focal element for the mass assignment* $\mathbf{m_s}$ *if*

$$m_s(A) > 0 \quad (3)$$

Example 1 *Let* $\mathbf{S} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots \mathbf{x}\}$. *Then a function* $\mathbf{m} : \mathcal{P}(\mathbf{S}) \mapsto [\mathbf{0}, \mathbf{1}]$ *defined by*

$$m(A) = \begin{cases} 0.6 & A = \{a, b\} \\ 0.4 & A = \{b, c\} \\ 0 & otherwise \end{cases}$$

is a MA on $\mathbf{S}$ *with focal elements* $\{\mathbf{a}, \mathbf{b}\}$ *and* $\{\mathbf{b}, \mathbf{c}\}$.

MAT and Probability

Based on MAT, the following relation between a (discrete) probability distribution (e.g. a normalized histogram) associated to elements of a sample space $\mathbf{S}$ and a given MA $\mathbf{m_s}$ on $\mathbf{S}$ can be derived

$$P_s(x) = \sum_{A \subseteq S, x \in A} P_A(x) \cdot m_s(A) \quad (4)$$

where $\mathbf{P_s(x)}$ is a probability distribution of $\mathbf{x}$ on $\mathbf{S}$, and $\mathbf{P_A(x)}$ is a probability distribution on a focal element $\mathbf{A}$ of a given mass assignment $\mathbf{m_s}$. For convenience, as in (Ralescu 97) we call $\mathbf{P_A(x)}$ a selection rule.

Example 2 *Let* $P_{\{a,b\}}(a) = P_{\{a,b\}}(b) = 0.5$ *and* $P_{\{b,c\}}(b) = P_{\{b,c\}}(c) = 0.5$ *be selection rules in Example 1 above. Then using (4), we calculate the probabilities of a, b, and c as:*

$$\begin{aligned} P_S(a) &= P_{\{a,b\}}(a) \times m(\{a,b\}) \\ &= 0.5 \times 0.6 = 0.3 \\ P_S(b) &= P_{\{a,b\}}(b) \times m(\{a,b\}) \\ &\quad + P_{\{b,c\}}(b) \times m(\{b,c\}) \\ &= 0.5 \times 0.6 + 0.5 \times 0.4 \\ &= 0.5 \\ P_S(c) &= P_{\{b,c\}}(c) \times m(\{b,c\}) \\ &= 0.5 \times 0.4 = 0.2 \end{aligned}$$

The selection rules in Example 2 are equally distributed and thus called *least prejudiced distribution*(LPD). On the other hand, a selection rule is called *most prejudiced distribution*(MPD) if it assigns its entire mass to only one element. For example, the selection rules $P_{\{a,b\}}(a) = 1.0$ *and* $P_{\{a,b\}}(b) = 0.0$ are MPD because **a** has its entire distribution. Obviously, there are many choices for selection rules. The flexibility of representation provided by PIP follows from these choices.

MAT and Fuzzy Sets

Let $S = \{x_1 \ldots x_n\}$ and $\mathbf{F} = \mathbf{x_1}/\mu_1 + \mathbf{x_2}/\mu_2 + \ldots + \mathbf{x_n}/\mu_n$ a fuzzy set on $\mathbf{S}$ specified through its membership function $\mu_F : S \longrightarrow [0, 1]$. We denote $\mu_i = \mu_F(x_i)$ and without loss of generality we assume

$$1 = \mu_1 \geq \mu_2 \geq \ldots \geq \mu_n > \mu_{n+1} = 0$$

Then, according to Baldwin (Baldwin 95), a mass assignment $\mathbf{m_f}$ with nested focal elements $\{x_1, \ldots, x_i\}$ where $i = 1, \ldots, n$ associated to $\mathbf{S}$ can be derived as follows:

$$m_f(A) = \begin{cases} \mu_i - \mu_{i+1} & \text{if } A = \{x_1, \ldots, x_i\} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Example 3 *Let* $S = \{x_1, x_2, x_3\}$ *and* $\mathbf{f}$ *be a fuzzy set defined as*

$$f = x_1/1 + x_2/0.7 + x_3/0.5$$

Then

$$\begin{aligned} m_f(\{x_1\}) &= 0.3 \\ m_f(\{x_1, x_2\}) &= 0.2 \\ m_f(\{x_1, x_2, x_3\}) &= 0.5 \end{aligned}$$

is a MA. A fuzzy set $\mathbf{f}$ *based on this relation is computed as follows:*

$$\begin{aligned} x_1 : \quad & \mu_1 &= 1 \\ x_2 : \quad & m_f(\{x_1\}) &= \mu_1 - \mu_2 = 0.3 \\ & &\Rightarrow \mu_2 = 0.7 \\ x_3 : \quad & m_f(\{x_1, x_2\}) &= \mu_2 - \mu_3 = 0.2 \\ & &\Rightarrow \mu_3 = 0.5 \end{aligned}$$

It can be seen that (5) can also be used to derive a normal fuzzy set, i.e. a fuzzy set whose maximum membership value is 1.0, from a given MA with nested focal elements as Example 3 illustrates.

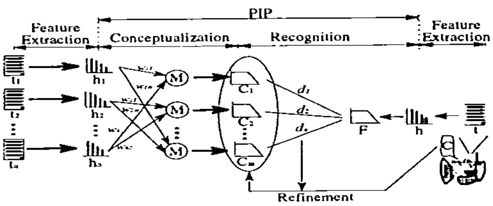

Figure 1: Overview of PIP Computational Model

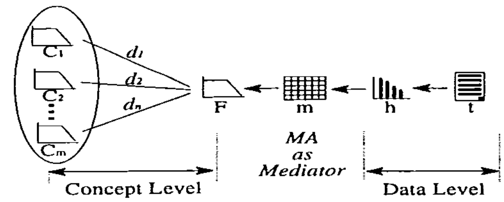

Figure 2: Recognition in PIP

Probability and Fuzzy Set

Using equations (4) and (5), we obtain the mapping between probability and a fuzzy set as in equation (6). Let $P_s(x_k)$ be a probability of a sample space $\mathbf{S}$, and $P_{A_i}(x_k)$ be a selection rule for $\mathbf{x_k}$ from the focal element $\mathbf{A_i} = \{\mathbf{x_1}, \ldots, \mathbf{x_i}\}, \mathbf{i} = \mathbf{1}, \ldots, \mathbf{n}$ of a MA. Then

$$P_s(x_k) = \sum_{i=k}^{n} P_{A_i}(x_k) \cdot (\mu_i - \mu_{i+1}) \qquad (6)$$

Operations of MAT

Operations on mass assignments are defined in a way compatible to set operations (Baldwin 95). They include the complement ($\bar{\ }$), meet ($\wedge$), and join ($\vee$). The probability distributions are redistributed over focal elements of MA as a result as either appropriate set operations such as intersection, union, and complement of the distributions of original mass assignments or a linear combination of original mass assignments orthogonal each other if the result of the set operations is empty.

Perceptual Information Processing (PIP)

As shown in Figure 1, PIP consists of three main tasks: conceptualization, recognition, and refinement. The conceptualization task acquires the user perception either by an aggregation of examples as training data sets or by concrete prototypes for each concept for each user, i.e. manual configurations. The recognition task determines results of intelligent information tasks such as categorization and recognition of matching degrees. The refinement task refines acquired concepts based on the result of determination and the user's evaluation.

Recognition

The steps of generating a fuzzy set from the data with the corresponding MA is as follows (See also Figure 2):

1. Generate a histogram by extracting features.
2. Normalize the histogram as relative frequencies.
3. Sort elements of the histogram in non-decreasing order of frequencies.
4. Convert the sorted histogram into a mass assignment with nested focal elements using (4). Choose appropriate selection rules if necessary. Otherwise, LPD is selected.
5. Convert the mass assignment to a fuzzy set using (5).

Once the fuzzy set is generated, a similarity degree taking values in $[0, 1]$ between this and a fuzzy set representing a concept is determined by using various definitions of the degree (Klir and Yuan 95) such as index of intersection (inclusion) (See Equation (11)).

It depends on the application whether the magnitude or rank of the matching degree are important or not. Categorization problems, for instance, do not consider magnitudes of the degrees important. Assignment of classes/concepts is determined in terms of the order of degrees that is, a data set is categorized as i-th category if

$$d(F|C_i) > d(F|C_j) \forall j \neq i \qquad (7)$$

where F is the fuzzy set generated from the data set, C is the fuzzy set representing user's perception for a category, and $d(F|C)$ is the degree of which the data set belongs to the category in terms of the matching degree of fuzzy sets.

Conceptualization

Conceptualization in PIP is a process to generate a fuzzy set for a concept $\mathbf{C}$ either from multiple data sets with weights in $[0, 1]$ or by user defined arbitrary features with weights in $[0, 1]$.

Let $\mathbf{F_C}$ denote the fuzzy set for concept $\mathbf{C}$ and $\mathbf{F_{user}}$ the fuzzy set for the same concept as defined by a user. Let $\mathbf{h}$ denote a histogram. For each example of the concept we will derive a histogram from which a fuzzy set $\mathbf{f(h)}$ will be obtained following the procedure described in the previous section. Typically conceptualization takes place after several concrete instances have been considered along with $\mathbf{F_{user}}$. Let $\mathbf{h_i}$, $i = 1 \ldots n$ be the histograms corresponding to such instances and $\mathbf{w_i}$, $i = 1, \ldots, n$ the weights corresponding to the $i^{\underline{th}}$ instance. Then the conceptualization can be expressed as

$$F_C = \mathcal{H}_{\underline{w}}(f(h_1), \ldots, f(h_n)) \oplus F_{user} \qquad (8)$$

Level	Functions
Histogram	Ave, Max, Min, Sum, Prod
MA	Join, Meet
Selection Rules	Selection Rule Generation
Fuzzy Set	AND, OR

Table 1: List of Aggregation

where $\mathcal{H}_{\underline{w}}$ denotes the aggregation of $f(h_i)$, $i = 1 \ldots n$, given the weight vector $\underline{w}$ and $\oplus$ denotes a set operation (e.g. union).

In the aggregation procedure, multiple data sets are converted into their histograms in the same way as in recognition in PIP. Then the aggregation is performed as part of the process of transforming histograms into MA's and then fuzzy sets using (4) and (5). The procedure uses either operators themselves, their combinations, or modification at certain level satisfying properties of aggregation operations described in reference (Klir and Yuan 95) such as histogram, MA's, and fuzzy sets shown in Table 1.

Refinement

Refinement in PIP is a process of adapting a concept using an arbitrary data set right after recognition process is attempted so as to improve the result at subsequent uses. The refinement takes as input two fuzzy sets, $\mathbf{F_C}$ the fuzzy set obtained at conceptualization and $\mathbf{F_D}$ a fuzzy set corresponding to a new data set. The result of the reinforcement, $\mathbf{F'_C}$ is defined as an aggregation $\mathcal{H}$ of $\mathbf{F_C}$ and $\mathbf{F_D}$, that is

$$F'_C = \mathcal{H}(F_C, F_D) \tag{9}$$

Possible aggregations of these sets include all the aggregation operations for conceptualization described in the previous section.

In order for the conceptualization results to agree it is necessary to use the same aggregation operator.

Adaptive Behavior

Description

In general, adaptive behavior (Staddon 89) is the process whereby a system's performance improves each time the system is being used. In other words, the more the system works, the better it becomes at the task for which it is intended. In PIP the recognition results reflect better and better a user's perception of the data. In engineering applications, neural networks are often considered as the typical solutions for the implementation of such process.

However, we believe, and this paper is a step towards demonstrating this, that PIP preserves some of the aspects of neural network systems, while at the same time offers a more transparent view of the underlying process. For problems in which human-system interactions are essential this transparency is expected to be of great advantage.

	D_1	D_2	D_3	D_4	D_5	D_6
a	4	2	0	0	0	0
b	3	3	2	1	0	0
c	2	0	0	0	0	0
d	1	0	0	0	1	2
e	0	1	0	0	0	0
f	0	0	1	3	1	2
g	0	1	0	2	2	0
h	0	0	0	2	0	1
i	0	0	0	0	1	0
j	0	0	0	0	0	0
total	10	7	3	8	5	5

Table 2: Data Sets

In case of PIP, recognition results improve each time a refinement takes place, which updates a recognized data set with the difference of the results given both from the system and the user.

The example in the following section illustrates this aspect. In this small and simple example, we use fuzzy set union operator such that

$$F'_C = F_C \cup F_{D_i} \tag{10}$$

as the aggregation, the index of intersection such that

$$d(F|C) = \frac{|F \cap C|}{|F|} \tag{11}$$

as the matching result for recognition process, and the sigma-count fuzzy cardinality, i.e. the sum of all membership values of a fuzzy set as the cardinality, for the sake of simplicity.

Example

Let D_1 through D_6 shown in Table 2 denote the data sets given as input to our system. Following an initial processing, fuzzy sets F_1 through F_6 shown in Table 4 are generated using the LPD selection rules from the corresponding mass assignments shown in Table 3.

Let us suppose that there are two categories namely C_1 and C_2 and further that data sets D_1, D_2, and D_3 belong to category C_1 and D_4, D_5, and D_6 belong to category C_2. Concept fuzzy set F_{C_1} (F_{C_2}) is consecutively refined in the following order:

1. Start with a empty fuzzy set F_{C_1} (F_{C_2}) and recognize all data sets for category C_1 (C_2)
2. Refine with data set D_1 (D_4) and recognize all data sets for category C_1 (C_2)
3. Refine with data set D_2 (D_5) and recognize all data sets for category C_1 (C_2)
4. Refine with data set D_3 (D_6) and recognize all data sets for category C_1 (C_2)

finement tasks, and regardless of the order should be the same when the same set of examples are used.

In order to satisfy this, aggregations of conceptualization should be identical to the one of refinement, and they must be symmetric. It is well known that aggregation operators of fuzzy sets are by definition symmetric (Klir and Yuan 95).

In this case, it is at the level of fuzzy set, and fuzzy set union operator that the aggregation satisfies this property. Similarly, join operator of mass assignment (at least for multiplication join) and summation for histogram satisfy this.

Improvement on Precision of Recognition

Adaptive behavior of animals consists in the improvement in results and stability of recognition task each time recognition is performed. Our simple, yet quite general, experiment illustrates a similar behavior of the system each time refinement is done. Therefore it suggests that PIP has adaptive behavior.

In conjunction with the ability to manage consistency between data and its higher level summarization (fuzzy sets) this adaptivity is an important advantage of the Perceptual Information Processing method. A further appealing feature from a theoretical point of view is the "single-framework" aspect of this method by contrast with other, hybrid approaches, such as combination of ANN and fuzzy sets which tend to be rather ad-hoc and complex and fail to keep such consistencies.

In future work we will show that such improvement occurs regardless of the choice of operators for aggregation, recognition, and refinement.

Interference of Adaptation

It is expected that the best adaptive behavior will cause improvement only for the refined concept. Yet, as already mentioned, because of the increase of membership values of elements common to both concepts, improvement for both concepts is detected.

Moreover, the matching degree becomes better as the concept is refined but it never becomes worse because of the linearly increasing characteristic of the operation. This does not have to be a bad feature as long as differentiation among of matching degree is clearly made for categorization. However, it becomes a problem once degrees themselves are important.

To avoid such interference of the refinement of one concept to the other concepts, fuzzy union operation could be replaced by one of the operations below:

$$F'_C = F_C \cup (F_D \cap \overline{F_C}) \quad (12)$$

$$F'_C = (\overline{F_D} \cap F_C) \cup (F_D \cap \overline{F_C}) \quad (13)$$

However, the use of these operations will cause a discrepancy between refinement and conceptualization result.

Conclusion

The adaptive behavior of PIP is illustrated by a simple example. Fuzzy union operator refines a concept identically to the concept fuzzy set generated by conceptualization task using the identical data sets and the aggregation operator at the same level, i.e. fuzzy set level. Further work will aim at understanding better the limits of this mechanism and at considering other, more general, operators.

References

J. F. Baldwin, T. P. Martin, B. W. Pilsworth, *FRIL-Fuzzy and Evidential Reasoning in Artificial Intelligence*, Research Studies Press, 1995.

A. Inoue, A. L. Ralescu, *Perceptual Text Categorization (PTC) :Application of Perceptual Information Processing (PIP)*, Proc. of 2nd Japan-Australia Joint Workshop on intelligent and evolutionary Systems, pp. 22-29, Nov 1998.

G. J. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, 1995.

A. L. Ralescu, *Quantitative Summarization of Numerical Data Using Mass Assignment Theory*, Unpublished Manuscript, May 1997.

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

J.E.R. Staddon, R.H. Ettinger, *Learning: An Introduction to the Principle of Adaptive Behavior*, Harcourt Brace Jovanovich, Inc., 1989.

L. A. Zadeh, *Fuzzy Logic = Computing with Words*, IEEE Transactions of Fuzzy Systems, Vol. 4, No. 2, pp. 103-111, May 1996.

L. A. Zadeh, *Fuzzy Sets*, Information and Control, Vol. 8, pp. 338-353, 1965.