Int. J. Appl. Math. Comput. Sci., 2000, Vol.10, No.4, 675–701

IMPLICATION-BASED NEURO-FUZZY ARCHITECTURES

DANUTA RUTKOWSKA*, ROBERT NOWICKI*

This paper presents connectionist multi-layer architectures of neuro-fuzzy systems based on various fuzzy implications. The well-known Mamdani approach (constructive) and the logical approach (destructive) are considered. Two kinds of architectures, a simpler and a more general one, are distinguished. Examples of application to classification and control problems are provided.

Keywords: neuro-fuzzy systems, fuzzy implications, fuzzy inference, Mamdani approach, logical approach, connectionist architectures

1. Introduction

A growing interest in the fusion of neural networks and fuzzy systems takes advantage of merits of the both methods and leads to various approaches of combining them into a form of neuro-fuzzy systems. Different neuro-fuzzy systems can be created using various types of fuzzy implications which correspond to fuzzy IF-THEN rules. Specific architectures of the systems can be found in (Rutkowska, 1997; Rutkowska *et al.*, 1997; Wang, 1994). However, they are confined to Mamdani's and Larsen's types of fuzzy inference, most often applied in fuzzy controllers. In this paper, we consider neuro-fuzzy systems based on the following implications: Kleene-Dienes, Łukasiewicz, Reichenbach, Zadeh, Willmott, Goguen, Gödel, Sharp, Fodor, Yager (Cordon *et al.*, 1997a; Cordon *et al.*, 1997b; Driankov *et al.*, 1993), and compare them with the well-known systems which employ Mamdani's or Larsen's rules of inference. The system architectures are connectionist and multi-layer, like those of artificial neural networks (Żurada, 1992). These systems can be trained using the idea incorporated into the back-propagation algorithm, commonly used to train neural networks. Applications of the systems to classification and control problems are also illustrated.

Two main approaches to the inference of fuzzy (or neuro-fuzzy) systems can be distinguished: the Mamdani (constructive) approach and the *logical* (destructive) approach (Czogała and Łęski, 1997; Czogała *et al.*, 1997a; Filev and Yager, 1995; Nowicki and Rutkowska, 1999; Rutkowska and Nowicki, 1999; Yager and Filev, 1994). The former is well-known and commonly used in approximate (fuzzy) reasoning. It is based on Mamdani's or Larsen's types of fuzzy inference. The latter is not so popular,

* Technical University of Częstochowa, Department of Computer Engineering, al. Armii Krajowej 36, 42–200 Częstochowa, Poland, e-mail: {drutko,rnowicki}@kik.pcz.czest.pl

however it has also been studied in the literature mentioned above. This paper refers mostly to the latter approach. It should be emphasised that although the names *constructive* and *destructive* have been introduced in (Yager and Filev, 1994) and used by other authors, the name *logical* seems more adequate, taking into account the implication which is employed in the fuzzy inference.

We consider NOCFS and OCFS neuro-fuzzy systems. The former stand for non-overlapping consequent fuzzy sets and the latter stand for overlapping consequent fuzzy sets, and both refer to the fuzzy sets in the consequent part of fuzzy IF-THEN rules. These two kinds of neuro-fuzzy systems were introduced in (Nowicki and Rutkowska, 2000a) and studied in (Rutkowska and Nowicki, 2000; Rutkowska *et al.*, 2000). The OCFS systems are extensions of the NOCFS ones. It is worth noticing that most of the neuro-fuzzy system research concern the NOCFS systems, which can be treated as special cases of the more general OCFS systems, presented in this paper.

The paper pertains to neuro-fuzzy architectures proposed in (Nowicki, 1999), and published in conference proceedings (Nowicki and Rutkowska, 1999; 2000a; 2000b; 2000c; Rutkowska and Nowicki, 1999; 2000; Rutkowska *et al.*, 1999; 2000). The results outlined there have been collected, supplemented, compared, and generalised here in the framework of the implication-based neuro-fuzzy architectures. More computer simulations illustrating the systems performance have been carried out. Conclusions drawn from these experiments, as well as other important remarks are included.

2. Implication-Based Fuzzy Systems

In this section, we present fuzzy logic systems (Czogała and Łęski, 2000; Driankov *et al.*, 1993; Rutkowska *et al.*, 1997; Wang, 1994). The inference process performed by them is described. Various fuzzy implications used by the inference engine, the main part of a fuzzy system, are considered. Different aggregation methods, with reference to the Mamdani (constructive) and logical (destructive) approaches, are depicted. The so-called 'pure' fuzzy systems and systems with a fuzzifier and a defuzzifier are distinguished.

2.1. Fuzzy Inference

Fuzzy systems are knowledge-based systems. They make a fuzzy inference based on a collection of fuzzy IF-THEN rules, called the rule base, expressed as follows:

$$R^k\colon\ \textbf{IF}\quad \mathbf{x}\ \text{is}\ A^k\quad \textbf{THEN}\quad y\ \text{is}\ B^k, \tag{1}$$

where $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbf{X} \subset \mathbb{R}^n$ and $y \in \mathbf{Y} \subset \mathbb{R}$ are linguistic variables, $A^k = A_1^k \times \cdots \times A_n^k$ and B^k are fuzzy sets characterised by membership functions $\mu_{A^k}(\mathbf{x})$, $\mu_{B^k}(y)$, respectively, $k = 1, \dots, N$. If $x_1, \dots, x_n$ are independent variables, the rule base (1) takes the form

$$R^k\colon\ \textbf{IF}\quad x_1\ \text{is}\ A_1^k\quad \text{AND} \dots \text{AND}\quad x_n\ \text{is}\ A_n^k\quad \textbf{THEN}\quad y\ \text{is}\ B^k. \tag{2}$$

Fuzzy IF-THEN rules (1) or (2) correspond to fuzzy relations $A^k \to B^k$, which are often called fuzzy implications, in spite of the fact that they are not always implications in a logical sense.

A fuzzy system inference determines a mapping from an input fuzzy set $A' \subset \mathbf{X}$ to an output fuzzy set $B' \subset \mathbf{Y}$, using the rule base (1) or (2). Each individual rule of the rule base conducts a mapping from the fuzzy set $A' \subset \mathbf{X}$ to a fuzzy set $\bar{B}^k \subset \mathbf{Y}$. Then the fuzzy set B' is obtained by an aggregation of the fuzzy sets $\bar{B}^k$, for $k = 1, \ldots, N$. This approach is called the *individual rule of inference* or FITA, which stands for *First Inference Then Aggregate*. Another approach is called the *composition based inference*, or FATI, which stands for *First Aggregate Then Inference* (Czogała and Łęski, 2000). The latter means that the aggregated rule base is used in the inference process.

The fuzzy sets $\bar{B}^k$, for $k = 1, \ldots, N$, inferred by the individual rules R^k, are characterised by the following membership functions:

$$\mu_{\bar{B}^k}(y) = \sup_{\mathbf{x} \in \mathbf{X}} \left\{ \mu_{A'}(\mathbf{x}) \overset{T}{\star} \mu_{A^k \to B^k}(\mathbf{x}, y) \right\}, \tag{3}$$

where $\mu_{A'}(\mathbf{x})$ and $\mu_{A^k \to B^k}(\mathbf{x}, y)$ are the membership functions of the input fuzzy set A' and the fuzzy relation $A^k \to B^k$, respectively, and $\overset{T}{\star}$ can be any T-norm operator. Formula (3) constitutes a well-known expression, called the *sup-star* composition, which corresponds to the compositional rule of inference, expressed as follows:

$$\bar{B}^k = A' \circ \left(A^k \to B^k\right). \tag{4}$$

It was introduced by Zadeh (1973).

If the input fuzzy set A' is a fuzzy singleton, which means that the membership function of this fuzzy set equals 1 for $\mathbf{x} = \bar{\mathbf{x}}$ and is zero for $\mathbf{x} \neq \bar{\mathbf{x}}$, where $\bar{\mathbf{x}} = [\bar{x}_1, \ldots, \bar{x}_n]^T \in \mathbf{X} \subset \mathbb{R}^n$, then (3) takes the following form:

$$\mu_{\bar{B}^k}(y) = \mu_{A^k \to B^k}(\bar{\mathbf{x}}, y). \tag{5}$$

2.2. Fuzzy Implications

As seen in Section 2.1, the result of a fuzzy system inference depends on the fuzzy relation $A^k \to B^k$. Most often the input fuzzy set is a fuzzy singleton. In this case a membership function of the output fuzzy set inferred by an individual IF-THEN rule equals a membership function of the fuzzy relation, for $\mathbf{x} = \bar{\mathbf{x}}$, according to (5). Thus, it seems to be very important what kind of fuzzy relation (fuzzy implication) is employed in a fuzzy system.

There are many fuzzy implication operators known in the literature (Cordon *et al.*, 1997a; 1997b). However, the ones most often applied in fuzzy systems, especially in fuzzy controllers, are Mamdani's and Larsen's types of inference, based on the minimum and product operators, respectively (see Table 1). As we have mentioned before, Mamdani and Larsen's implications are not implications in a logical sense. They should be interpreted rather as a conjunction of antecedent and consequent parts of fuzzy IF-THEN rules. They correspond to the T-norm operators (minimum and product). Other implications, listed in Table 1, can be interpreted in the spirit of classical logical implications, but they are expressed by more sophisticated formulae than Mamdani and Larsen's ones. The Kleene-Dienes implication, also named the Dienes-Rescher, Boolean or binary implication, represents a straighforward fuzzy interpretation of the implication in classical logic. Other fuzzy implications named after Łukasiewicz, Reichenbach, Fodor, Sharp, Goguen, Gödel, Yager, Zadeh, Willmott, are inserted in this table as examples of genuine fuzzy implications.

Table 1. Fuzzy implications.

Name	$\mu_{A^k \to B^k}(\mathbf{x}, y)$
Mamdani	$\min\left[\mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)\right]$
Larsen	$\mu_{A^k}(\mathbf{x}) \cdot \mu_{B^k}(y)$
Kleene-Dienes	$\max\left[1 - \mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)\right]$
Łukasiewicz	$\min\left[1, 1 - \mu_{A^k}(\mathbf{x}) + \mu_{B^k}(y)\right]$
Reichenbach	$1 - \mu_{A^k}(\mathbf{x}) + \mu_{A^k}(\mathbf{x}) \cdot \mu_{B^k}(y)$
Fodor	$\begin{cases} 1 & \text{if} \quad \mu_{A^k}(\mathbf{x}) \le \mu_{B^k}(y) \\ \max\left[1 - \mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)\right] & \text{if} \quad \mu_{A^k}(\mathbf{x}) > \mu_{B^k}(y) \end{cases}$
Sharp	$\begin{cases} 1 & \text{if} \quad \mu_{A^k}(\mathbf{x}) \le \mu_{B^k}(y) \\ 0 & \text{if} \quad \mu_{A^k}(\mathbf{x}) > \mu_{B^k}(y) \end{cases}$
Goguen	$\begin{cases} 1 & \text{if} \quad \mu_{A^k}(\mathbf{x}) = 0 \\ \min\left[1, \frac{\mu_{B^k}(y)}{\mu_{A^k}(\mathbf{x})}\right] & \text{if} \quad \mu_{A^k}(\mathbf{x}) > 0 \end{cases}$
Gödel	$\begin{cases} 1 & \text{if} \quad \mu_{A^k}(\mathbf{x}) \le \mu_{B^k}(y) \\ \mu_{B^k}(y) & \text{if} \quad \mu_{A^k}(\mathbf{x}) > \mu_{B^k}(y) \end{cases}$
Yager	$\begin{cases} 1 & \text{if} \quad \mu_{A^k}(\mathbf{x}) = 0 \\ \mu_{B^k}(y)^{\mu_{A^k}(\mathbf{x})} & \text{if} \quad \mu_{A^k}(\mathbf{x}) > 0 \end{cases}$
Zadeh	$\max\left[\min[\mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)], 1 - \mu_{A^k}(\mathbf{x})\right]$
Willmott	$\min\Big[\max[1 - \mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)], \max\left[\mu_{A^k}(\mathbf{x}), 1 - \mu_{B^k}(y), \min[1 - \mu_{A^k}(\mathbf{x}), \mu_{B^k}(y)]\right]\Big]$

2.3. Aggregation Methods

Usually, when Mamdani's or Larsen's type of inference is applied, a union operation (S-norm) is employed as the aggregation. It is called Mamdani's combination. If we apply the genuine fuzzy implications (see Table 1), the intersection (T-norm) is a reasonable operation for the aggregation. This method is called the Gödel combination (Czogała and Łęski, 2000). The Mamdani combination corresponds to Mamdani's (constructive) approach and the Gödel combination corresponds to the logical (destructive) approach to fuzzy inference.

With Mamdani's approach the aggregated output fuzzy set $B' \subset \mathbf{Y}$ is determined by the formulae:

$$B' = \bigcup_{k=1}^{N} \bar{B}^k \tag{6}$$

and

$$\mu_{B'}(y) = \underset{k=1}{\overset{N}{S}} \mu_{\bar{B}^k}(y), \tag{7}$$

where S is an S-norm operator (a generalized form for more than two arguments), usually chosen as the maximum operator.

Under the logical approach, the aggregation is realised as follows:

$$B' = \bigcap_{k=1}^{N} \bar{B}^k \tag{8}$$

and

$$\mu_{B'}(y) = \underset{k=1}{\overset{N}{T}} \mu_{\bar{B}^k}(y), \tag{9}$$

where T is a T-norm operator (a generalized form for more than two arguments), usually chosen as the minimum or the product operator.

This approach corresponds to the FITA type of inference. The fuzzy sets inferred by individual IF-THEN rules are aggregated. When the FATI type of inference is applied, individual fuzzy IF-THEN rules are aggregated first, which means that the output fuzzy set B' is inferred, according to the compositional rule of inference, in the following way:

$$B' = A' \circ \mathfrak{R}, \tag{10}$$

where the method of the rule aggregation can be realised (under Mamdani's approach) by the union operation:

$$\mathfrak{R} = \bigcup_{k=1}^{N} R^k. \tag{11}$$

In this case, the membership function of the fuzzy set B' is expressed as follows:

$$\mu_{B'}(y) = \sup_{x \in \mathbf{X}} \left\{ \mu_{A'}(\mathbf{x}) \overset{T}{\star} \max_{1 \le k \le N} \mu_{A^k \to B^k}(\mathbf{x}, y) \right\}. \tag{12}$$

It is easy to show (Lee, 1990, Rutkowska *et al.*, 1997) that in Mamdani's approach (Mamdani's or Larsen's types of inference) the FITA and FATI infer the same result (the output fuzzy set B').

2.4. Fuzzy Systems with a Fuzzifier and a Defuzzifier

The so-called 'pure' fuzzy systems perform the inference, according to formula (3), based on fuzzy implications (fuzzy relations) which correspond to IF-THEN rules in the form (1) or (2). The inference with aggregation, under the FITA or FATI approaches, determines an output fuzzy set B' given an input fuzzy set A'.

In order to use fuzzy systems in the applications where inputs and outputs are not fuzzy sets but real-valued variables, a fuzzifier and a defuzzifier is added to the system. The fuzzifier maps crisp points in $\mathbf{X}$ to fuzzy sets in $\mathbf{X}$. As mentioned in Section 2.1, the most commonly employed fuzzifier is the singleton fuzzifier. The defuzzifier maps fuzzy sets in $\mathbf{Y}$ to crisp points in $\mathbf{Y}$. There are many defuzzification methods (Driankov *et al.*, 1993). In this paper, we apply the method called the *centre of area* (COA). The following formula gives a discrete version of this method:

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \mu_{B'}(\bar{y}^k)}{\sum_{k=1}^{N} \mu_{B'}(\bar{y}^k)}, \tag{13}$$

where $\bar{y}$ is the crisp output value of the system, and $\bar{y}^k$, for $k = 1, \ldots, N$, are the points in $\mathbf{Y}$ such that

$$\mu_{B^k}(\bar{y}^k) = \max_{y} \{\mu_{B^k}(y)\}. \tag{14}$$

We usually assume that

$$\mu_{B^k}(\bar{y}^k) = 1. \tag{15}$$

The points $\bar{y}^k$ are called the *centres* of the membership functions $\mu_{B^k}(y)$, $k = 1, \ldots, N$.

As a matter of fact, (13) is a special case of the discrete version of the COA defuzzification method, where the discrete points in $\mathbf{Y}$ are the centres of the membership functions of fuzzy sets B^k. It is easy to show that this case of the COA is the same as the *centre average* (CA) defuzzification method:

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \mu_{\bar{B}^k}(\bar{y}^k)}{\sum_{k=1}^{N} \mu_{\bar{B}^k}(\bar{y}^k)} \tag{16}$$

for $k = 1, \ldots, N$, if Mamdani's approach is employed and the aggregation is defined by (6), (7) with the maximum operator as the S-norm.

3. Neuro-Fuzzy Architectures

Fuzzy systems with fuzzifiers and defuzzifier, described in Section 2 can be represented in the form of connectionist multi-layer architectures, similar to artificial neural networks (Żurada, 1992). This kind of fuzzy system representations, called the neuro-fuzzy architectures, is very helpful from the learning point of view. Neuro-fuzzy systems in the form of multi-layer networks can be trained with the use of algorithms, analogous to the well-known back-propagation, commonly applied to neural networks. The idea of these algorithms comes from the steepest-descent optimisation technique. Special software which realises this method can be employed in order to train multilayer neural networks or neuro-fuzzy architectures. An example of this software is the FLiNN programme (Piliński, 1997a). Therefore, various neuro-fuzzy architectures are proposed in this section.

3.1. General Multi-Layer Architectures

Let us consider a fuzzy system, introduced in Section 2, with a singleton fuzzifier and a discrete form of the COA defuzzifier, given by (13), which depicts a crisp output of the system while the crisp input is $\bar{\mathbf{x}}$. For the singleton fuzzifier, (5) describes the inference performed by a single rule R^k in the form (1) or (2), $k = 1, \dots, N$. The rule corresponds to the fuzzy relation (implication) $A^k \to B^k$. Thus, from (13), (7), (9) and (5), we obtain the following description of the fuzzy system:

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \lambda_k \left(\bar{\mathbf{x}}, \bar{y}^k\right)}{\sum_{k=1}^{N} \lambda_k \left(\bar{\mathbf{x}}, \bar{y}^k\right)}, \tag{17}$$

where

$$\lambda_k \left(\bar{\mathbf{x}}, \bar{y}^k\right) = \underset{j=1}{\overset{N}{S}} \mu_{A^j \to B^j} \left(\bar{\mathbf{x}}, \bar{y}^k\right) \tag{18}$$

in the case of Mamdani's (constructive) approach, and

$$\lambda_k \left(\bar{\mathbf{x}}, \bar{y}^k\right) = \underset{j=1}{\overset{N}{T}} \mu_{A^j \to B^j} \left(\bar{\mathbf{x}}, \bar{y}^k\right) \tag{19}$$

in the case of the logical (destructive) approach.

It is easy to notice that based on (17)–(19) we can represent the fuzzy system described by these expressions in the form of the multi-layer, connectionist networks depicted in Figs. 1 and 2, respectively. Figure 1 illustrates the neuro-fuzzy architecture in the case of Mamdani's approach and Fig. 2 shows a similar architecture in the case of the logical approach. These two architectures differ in the third layer, where there are elements which implement an S-norm operator in the architecture presented in Fig. 1 and a T-norm operator in the architecture depicted in Fig. 2. This layer is called the *aggregation layer*. The other layers are the same in both the

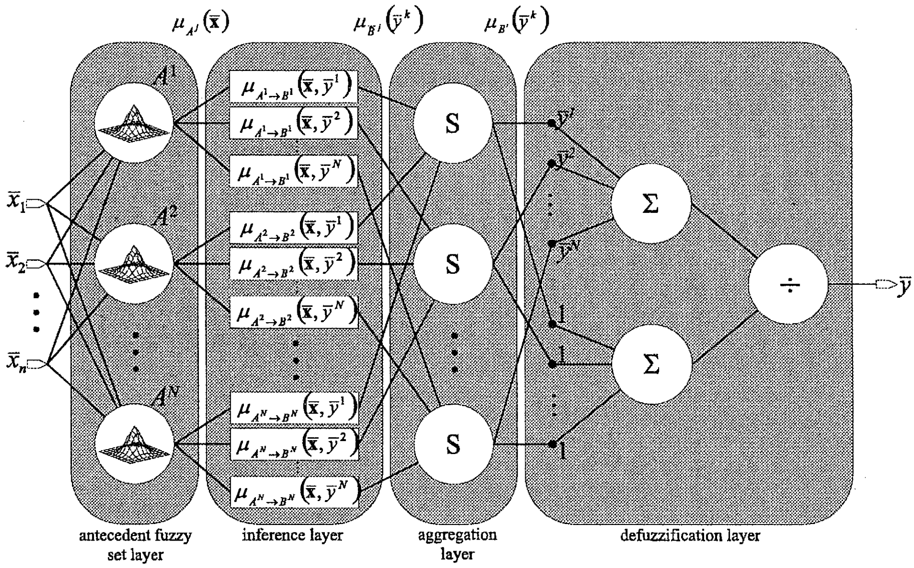

Fig. 1. General architecture of implication-based neuro-fuzzy systems in the case of Mamdani's (constructive) approach.

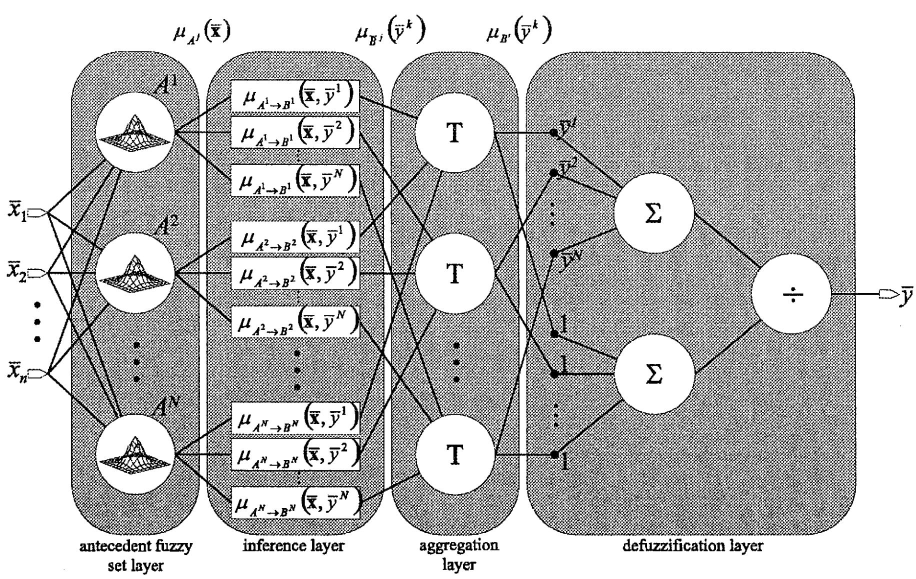

Fig. 2. General architecture of implication-based neuro-fuzzy systems in the case of the logical (destructive) approach.

neuro-fuzzy architectures. The first one contains the elements which realise the membership functions $\mu_{A^k}(\mathbf{x})$ of the fuzzy sets A^k, $k = 1, \ldots, N$. These fuzzy sets belong to antecedent parts of IF-THEN rules (1) or (2). Therefore, this layer is called the *antecedent fuzzy set layer.* The membership functions of these fuzzy sets can be e.g. Gaussian membership functions, as shown in Figs. 1 and 2. The second layer is called the *inference layer.* It performs the inference based on the fuzzy relations (implications) $A^k \to B^k$ which correspond to the IF-THEN rules. The last layer realises the defuzzification, according to (13). Thus, this layer is called the *defuzzification layer.* There are two typical neurons (Żurada, 1992) in this layer and the element which performs division. The weights of the first neuron are the centres of the membership functions $\mu_{B^k}(y)$, cf. (14). The weights of the second neuron equal 1. The centres of the membership functions $\mu_{B^k}(y)$, as well as the centres and widths of the membership functions $\mu_{A^k}(\mathbf{x})$, can be trained in much the same way as weights in artificial neural networks. Figures 1 and 2 present general architectures of implication-based neuro-fuzzy systems.

3.2. NOCFS Architectures in Mamdani's Approach

The inference layer in the general neuro-fuzzy architectures, introduced in Section 3.1, contains elements which perform fuzzy relations (implications). For a particular type of fuzzy relation $A^k \to B^k$, $k = 1, \ldots, N$, we obtain a special case of the neuro-fuzzy architecture which corresponds to this relation. Examples of the fuzzy implications are listed in Table 1.

In order to find a special case of the system description (17) for a particular fuzzy relation, let us start from (18) and (19). Let us notice that the S-norm in (18) can be expressed as follows:

$$\mathop{S}_{j=1}^{N} \mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = S\left\{ \mu_{A^k \to B^k}\left(\bar{\mathbf{x}}, \bar{y}^k\right), \mathop{S}_{\substack{j=1 \\ j \neq k}}^{N} \mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) \right\}. \tag{20}$$

Let us consider Mamdani's (constructive) approach to fuzzy inference with Mamdani's (minimum operator) or Larsen's (product operator) types of inference, cf. the first two rows of Table 1. In the case of Mamdani's implication (relation), we have

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \min\left[\mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)\right] \tag{21}$$

and in the case of Larsen's implication (relation), we get

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \mu_{A^j}(\bar{\mathbf{x}})\mu_{B^j}(\bar{y}^k) \tag{22}$$

for $j, k = 1, \ldots, N$.

Let us assume that (15) is fulfilled. This means that $\mu_{B^j}(\bar{y}^k) = 1$ for $j = k$. Therefore, from (18) and (20), we obtain

$$\lambda_k\left(\bar{\mathbf{x}}, \bar{y}^k\right) = S\left\{ \mu_{A^k}\left(\bar{\mathbf{x}}\right), \mathop{S}_{\substack{j=1 \\ j \neq k}}^{N} \mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) \right\}, \tag{23}$$

where $\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k)$ is defined by (21) and (22) for Mamdani's and Larsen's implications, respectively.

Let us notice that if

$$\mu_{B^j}(\bar{y}^k) = 0 \quad \text{for} \quad j \neq k, \tag{24}$$

then (23) takes the form

$$\lambda_k(\bar{\mathbf{x}}, \bar{y}^k) = \mu_{A^k}(\bar{\mathbf{x}}) \tag{25}$$

and from (17) we obtain

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \mu_{A^k}(\bar{\mathbf{x}})}{\sum_{k=1}^{N} \mu_{A^k}(\bar{\mathbf{x}})}. \tag{26}$$

In Section 2.4 we have seen that in Mamdani's approach with the maximum operator as the S-norm aggregation, the COA defuzzification (13) becomes a CA defuzzification method (16). Thus, it is worth showing that from the CA defuzzification, defined by (16), and from (5), (21) and (22), we obtain the same result as (26).

Formula (26) describes the fuzzy system based on Mamdani's or Larsen's implications if the assumptions (15) and (24) are fulfilled. The former is easy to fulfil. It requires the maximal values of the membership functions $\mu_{B^k}(y)$, $k = 1, \ldots, N$, to be equal to 1, which means that fuzzy sets B^k must be *normal.* The latter assumption requires that B^k's cannot overlap. Figure 3 shows non-overlapping fuzzy sets. Of course, a similar illustration can be given, e.g. for triangular membership functions. Since the fuzzy sets B^k, $k = 1, \ldots, N$, belong to the consequent parts of the IF-THEN rules (1) or (2), we call them *non-overlapping consequent fuzzy sets* (NOCFS).

Fig. 3. Non-overlapping consequent fuzzy sets.

The neuro-fuzzy architecture which corresponds to the system description (26) is depicted in Fig. 4. It is worth noticing that this architecture can be treated as a normalized version of the RBF (*radial basis function*) network (Moody and Darken, 1989).

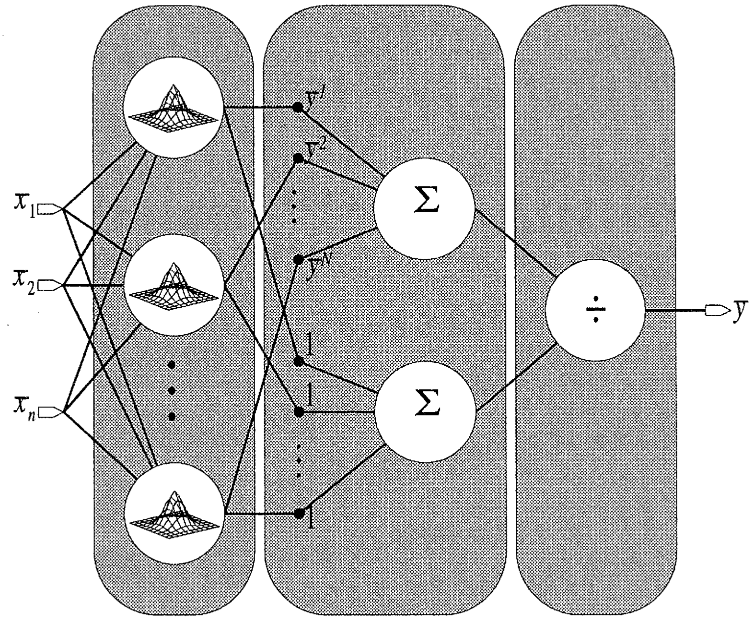

Fig. 4. NOCFS architecture based on Mamdani's or Larsen's implications.

3.3. NOCFS Architectures in the Logical Approach

In this section we present NOCFS architectures of the systems based on fuzzy implications used in the logical (destructive) approach to fuzzy inference. We consider implication-based neuro-fuzzy systems which employ other fuzzy implications, listed in Table 1, except the Mamdani and Larsen fuzzy relations. We assume that the conditions (15) and (24) are fulfilled and hence we apply non-overlapping consequent fuzzy sets (illustrated in Fig. 3).

In much the same way as in Section 3.2, the T-norm in (19) can be written as

$$\mathop{T}_{j=1}^{N} \mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = T\left\{ \mu_{A^k \to B^k}\left(\bar{\mathbf{x}}, \bar{y}^k\right), \mathop{T}_{\substack{j=1 \\ j \neq k}}^{N} \mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) \right\}. \tag{27}$$

For the Kleene-Dienes implication, we get

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \max\left[1 - \mu_{A^j}\left(\bar{\mathbf{x}}\right), \mu_{B^j}\left(\bar{y}^k\right)\right]. \tag{28}$$

For the Łukasiewicz implication, we have

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \min\left[1,\ 1 - \mu_{A^j}\left(\bar{\mathbf{x}}\right) + \mu_{B^j}\left(\bar{y}^k\right)\right]. \tag{29}$$

For the Reichenbach implication, it follows that

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = 1 - \mu_{A^j}(\bar{\mathbf{x}}) + \mu_{A^j}(\bar{\mathbf{x}})\mu_{B^j}(\bar{y}^k). \tag{30}$$

We can easily write similar expressions for other implications listed in Table 1.

Let us notice that for the Kleene-Dienes, Łukasiewicz, Reichenbach and Fodor implication, whenever (15) is fulfilled, then we get $\mu_{A^k \to B^k}(\bar{\mathbf{x}}, \bar{y}^k) = 1$. If (24) is fulfilled, then (28)–(30) become $\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k) = 1 - \mu_{A^j}(\bar{\mathbf{x}})$. The same result is obtained for the Fodor implication, since in this case (cf. Table 1) we have

$$\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k) = \begin{cases} 1 & \text{if } \mu_{A^j}(\bar{\mathbf{x}}) = 0, \\ 1 - \mu_{A^j}(\bar{\mathbf{x}}) & \text{if } \mu_{A^j}(\bar{\mathbf{x}}) > 0. \end{cases}$$

Therefore, from (17), (19) and (27), we get the following formula which describes the NOCFS systems based on the Kleene-Dienes, Łukasiewicz, Reichenbach and Fodor implications:

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \mathop{T}_{\substack{j=1 \\ j \neq k}}^{N} \left(1 - \mu_{A^j}(\bar{\mathbf{x}})\right)}{\sum_{k=1}^{N} \mathop{T}_{\substack{j=1 \\ j \neq k}}^{N} \left(1 - \mu_{A^j}(\bar{\mathbf{x}})\right)}. \tag{31}$$

Figure 5 presents the NOCFS architecture of the systems described by (31).

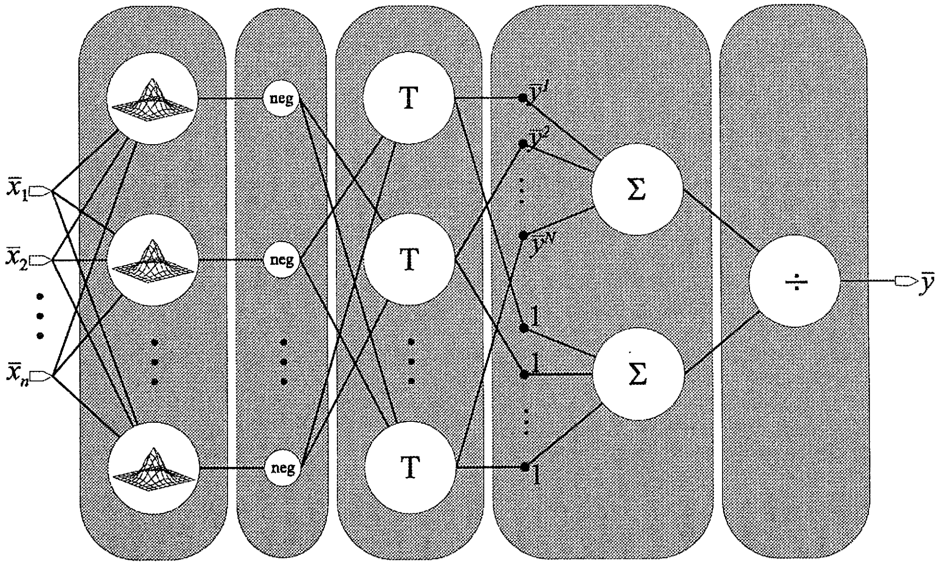

Fig. 5. NOCFS architecture based on binary, Łukasiewicz or stochastic implications.

Similarly, we can obtain formulae which describe the NOCFS systems based on other fuzzy implications.

Let us notice that for the Sharp, Goguen, Gödel and Yager implications (cf. Table 1), if the assumption (15) is fulfilled, then $\mu_{A^k \to B^k}(\bar{\mathbf{x}}, \bar{y}^k) = 1$. If the assumption (24) is fulfilled, then

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \begin{cases} 1 & \text{if } \mu_{A^j}\left(\bar{\mathbf{x}}\right) = 0, \\ 0 & \text{if } \mu_{A^j}\left(\bar{\mathbf{x}}\right) > 0. \end{cases} \tag{32}$$

From (17), (19), (27) and (32), we obtain the following formula which describes the NOCFS systems based on the Sharp, Goguen, Gödel or Yager implications:

$$\bar{y} = \frac{\sum\limits_{k=1}^{N} \bar{y}^k \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \delta\left(\mu_{A^j}\left(\bar{\mathbf{x}}\right)\right)}{\sum\limits_{k=1}^{N} \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \delta\left(\mu_{A^j}\left(\bar{\mathbf{x}}\right)\right)}, \tag{33}$$

where

$$\delta(a) = \begin{cases} 1 & \text{if } a = 0, \\ 0 & \text{if } a > 0, \end{cases} \tag{34}$$

for $a \in [0,1]$.

Figure 6 presents the connectionist architecture of the system described by (33). The multi-layer architectures illustrated in Figs. 5 and 6 differ in the second layer. In Fig. 5 this layer contains *neg* elements which realize the complement operations of fuzzy sets A^j, $j = 1, \ldots, N$, defined by $1 - \mu_{A^j}(\bar{\mathbf{x}})$, where μ_{A^j} is the membership function of the fuzzy set A^j. In Fig. 6 the second layer contains the elements which perform the operation defined by (34). The former is the architecture of the NOCFS neuro-fuzzy systems based on the Kleene-Dienes, Łukasiewicz or Reichenbach implications. The latter is the architecture of the NOCFS neuro-fuzzy systems based on the Sharp, Goguen, Gödel or Yager implications.

Now, let us consider Zadeh and Wilmott's fuzzy implications (cf. Table 1). In this case, if the assumption (15) is fulfilled, it is easy to show that $\mu_{A^k \to B^k}(\bar{\mathbf{x}}, \bar{y}^k) = \max[\mu_{A^k}(\bar{\mathbf{x}}), 1 - \mu_{A^k}(\bar{\mathbf{x}})]$. If the assumption (24) is fulfilled, then $\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k) = 1 - \mu_{A^j}(\bar{\mathbf{x}})$. Therefore, from (17), (19) and (27), we obtain the following formula which describes the NOCFS systems based on Zadeh's or Wilmott's implication:

$$\bar{y} = \frac{\sum\limits_{k=1}^{N} \bar{y}^k T\left\{ \max\left(\mu_{A^k}(\bar{\mathbf{x}}), 1 - \mu_{A^k}(\bar{\mathbf{x}})\right), \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \left(1 - \mu_{A^j}(\bar{\mathbf{x}})\right) \right\}}{\sum\limits_{k=1}^{N} T\left\{ \max\left(\mu_{A^k}(\bar{\mathbf{x}}), 1 - \mu_{A^k}(\bar{\mathbf{x}})\right), \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \left(1 - \mu_{A^j}(\bar{\mathbf{x}})\right) \right\}}. \tag{35}$$

Figure 7 presents the multi-layer, connectionist architectures of the NOCFS neuro-fuzzy systems based on Zadeh's or Wilmott's implication, described by (35).

The second layer in this architecture is the same as the second layer in the neuro-fuzzy architecture illustrated in Fig. 5. Another layer (the third one), with elements which perform the maximum operation, is added to this architecture.

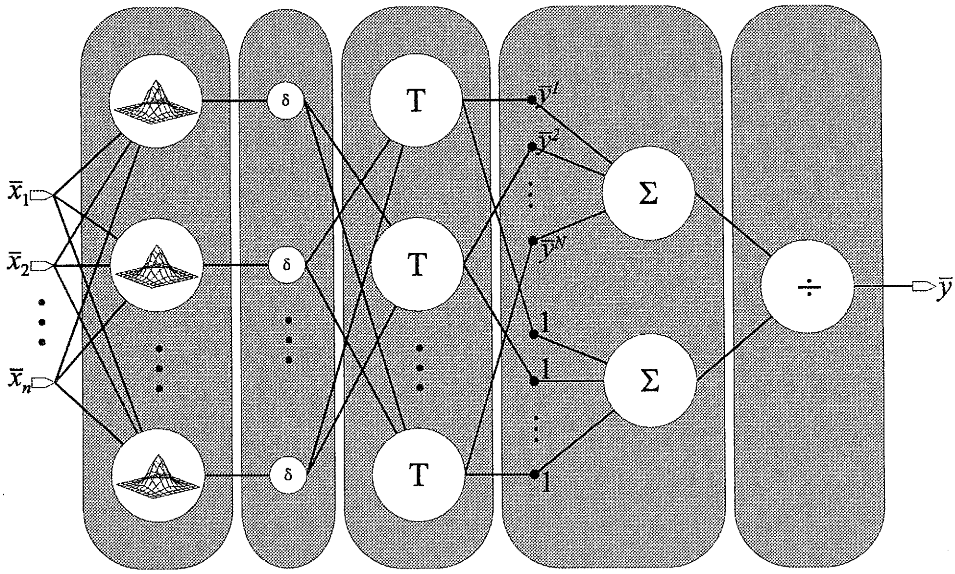

Fig. 6. NOCFS architecture based on the Sharp, Goguen, Godel or Yager implications.

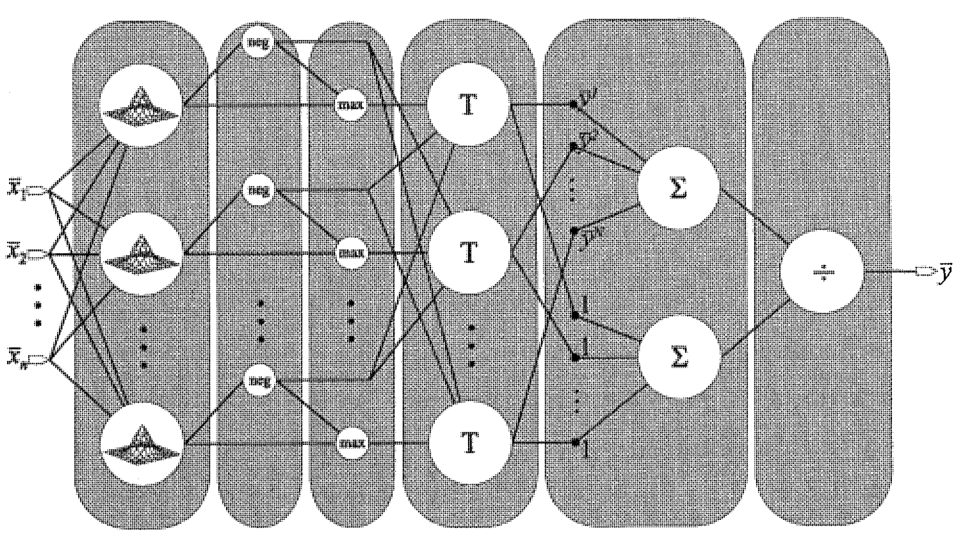

Fig. 7. NOCFS architecture based on Zadeh's or Willmott's implication.

3.4. OCFS Architectures in Mamdani's Approach

In Section 3.2 the neuro-fuzzy architectures with NOCFS have been presented. We can apply these architectures if the assumptions (15) and (24) are fulfilled, especially the latter. Now let us assume that (24) is not fulfilled, which means that we allow the consequent fuzzy sets B^k, $k = 1, \dots, N$, to overlap. Figure 8 illustrates the overlapping fuzzy sets. Analogously to the NOCFS, we call the fuzzy sets B^k which can overlap the *overlapping consequent fuzzy sets* (OCFS).

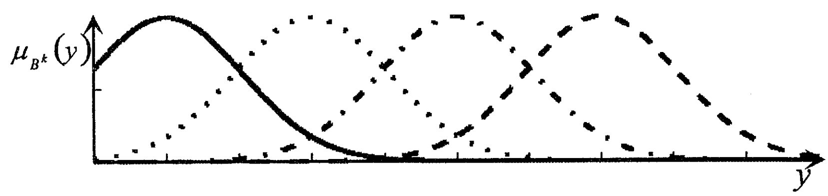

Fig. 8. Overlapping consequent fuzzy sets.

As in Section 3.2, let us consider Mamdani's (constructive) approach with Mamdani's (minimum operator) or Larsen's (product operator) types of inference. If the assumption (24) is not fulfilled, the system is described by (17) and (23), where $\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k)$ is defined by (21) or (22), respectively, for Mamdani's or Larsen's types of inference.

Write

$$p_{j,k} = \mu_{B^j}\left(\bar{y}^k\right). \tag{36}$$

Thus, the system description is expressed by (17), where

$$\lambda_k\left(\bar{\mathbf{x}}, \bar{y}^k\right) = S\left\{\mu_{A^k}(\bar{\mathbf{x}}), \underset{\substack{j=1\\ j\neq k}}{\overset{N}{S}} \min\left[\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}\right]\right\} \tag{37}$$

for the OCFS system based on Mamdani's implication, and

$$\lambda_k\left(\bar{\mathbf{x}}, \bar{y}^k\right) = S\left\{\mu_{A^k}(\bar{\mathbf{x}}), \underset{\substack{j=1\\ j\neq k}}{\overset{N}{S}} \mu_{A^j}(\bar{\mathbf{x}})p_{j,k}\right\} \tag{38}$$

for the OCFS system based on Larsen's implication.

Figures 9 and 10 present neuro-fuzzy architectures which correspond to the system descriptions (17) with (37) and (17) with (38), respectively.

It is easy to notice that if $p_{j,k} = 0$, $j, k = 1, \dots, N$, then (37) and (38) take the form of the expression (25) and in this case the neuro-fuzzy system is described by (26). Thus, the OCFS system reduces to the NOCF system presented in Section 3.2 and illustrated in Fig. 4.

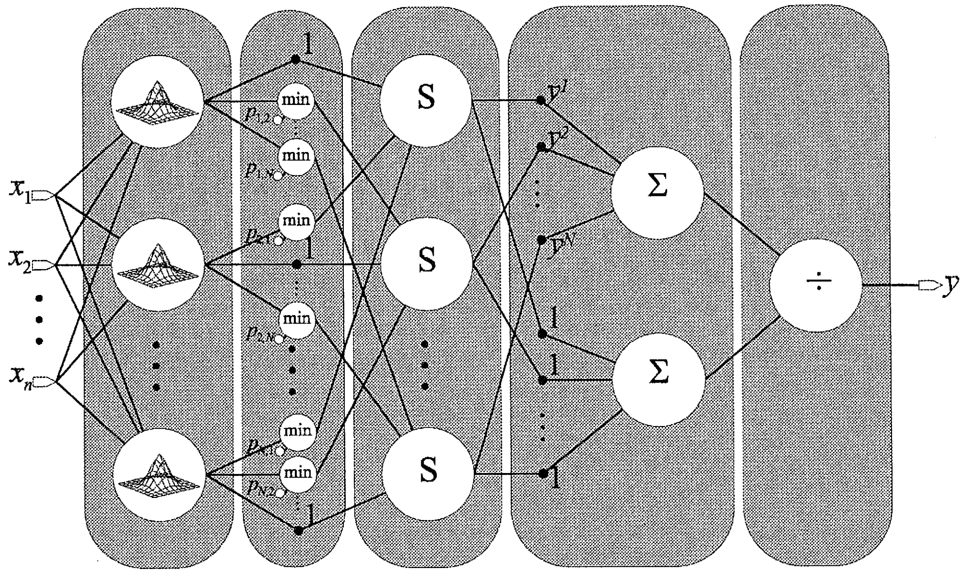

Fig. 9. OCFS architecture based on Mamdani's implication.

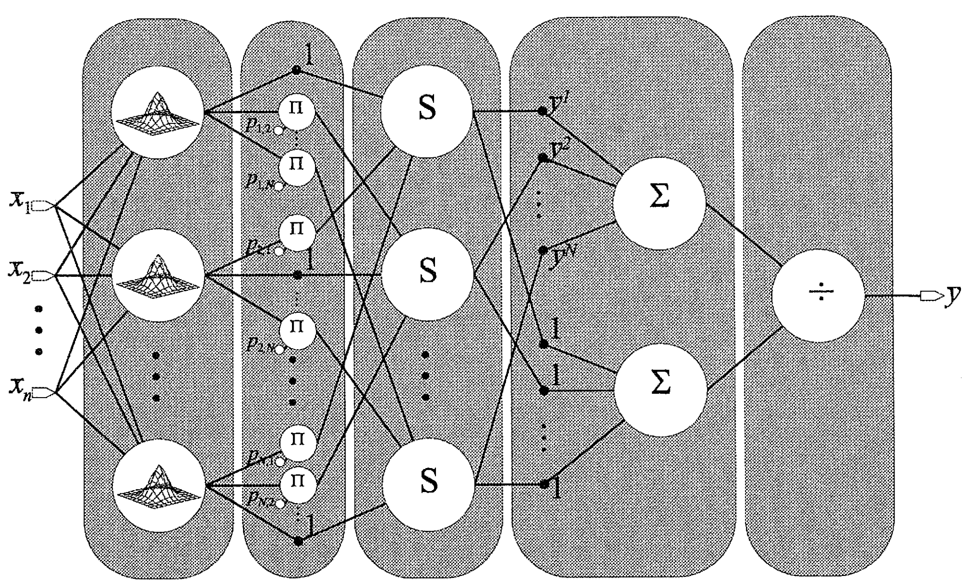

Fig. 10. OCFS architecture based on Larsen's implication.

3.5. OCFS Architectures in the Logical Approach

In this section, we present OCFS systems based on the implications listed in Table 1, but not Mamdani's and Larsen's types of inference. Similarly to Section 3.4, let us assume that the condition (24) is not fulfilled. The neuro-fuzzy systems are described by formula (17) with (19) and (27). As we noticed in Section 3.4, $\mu_{A^k \to B^k}(\bar{\mathbf{x}}, \bar{y}^k) = 1$ for the Kleene-Dienes, Łukasiewicz, Reichenbach and Fodor implications, as well as in the case of Sharp, Goguen, Gödel and Yager implications. Therefore, for these implications, from (19) and (27), we have

$$\lambda_k\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right). \tag{39}$$

It is easy to show, as in Section 3.3, that for Zadeh and Wilmott's fuzzy implications $\mu_{A^k \to B^k}(\bar{\mathbf{x}}, \bar{y}^k) = \max[\mu_{A^k}(\bar{\mathbf{x}}), 1 - \mu_{A^k}(\bar{\mathbf{x}})]$. Therefore, in this case, from (19) and (27), we obtain

$$\lambda_k\left(\bar{\mathbf{x}}, \bar{y}^k\right) = T\left\{ \max\left[\mu_{A^k}\left(\bar{\mathbf{x}}\right), 1 - \mu_{A^k}\left(\bar{\mathbf{x}}\right)\right], \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) \right\}. \tag{40}$$

From (17), (39), (28) and (36), we get the following description of the OCFS neuro-fuzzy system based on the Kleene-Dienes implication:

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \max\left[1 - \mu_{A^j}\left(\bar{\mathbf{x}}\right), p_{j,k}\right]}{\sum_{k=1}^{N} \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \max\left[1 - \mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}\right]}. \tag{41}$$

From (17), (39), (29) and (36), we have the following description of the OCFS neuro-fuzzy system based on the Łukasiewicz implication:

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \min\left[1, 1 - \mu_{A^j}(\bar{\mathbf{x}}) + p_{j,k}\right]}{\sum_{k=1}^{N} \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \min\left[1, 1 - \mu_{A^j}(\bar{\mathbf{x}}) + p_{j,k}\right]}. \tag{42}$$

From (17), (39), (30) and (36), we get the following description of the OCFS neuro-fuzzy system based on the Reichenbach implication:

$$\bar{y} = \frac{\sum_{k=1}^{N} \bar{y}^k \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \left(1 - \mu_{A^j}(\bar{\mathbf{x}}) + \mu_{A^j}(\bar{\mathbf{x}}) p_{j,k}\right)}{\sum_{k=1}^{N} \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \left(1 - \mu_{A^j}(\bar{\mathbf{x}}) + \mu_{A^j}(\bar{\mathbf{x}}) p_{j,k}\right)}. \tag{43}$$

The description of the OCFS system based on the Fodor implication can be obtained from (17), (39), (36) and the following expression (cf. Table 1):

$$\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k) = \begin{cases} 1 & \text{if } \mu_{A^j}(\bar{\mathbf{x}}) \leq \mu_{B^j}(\bar{y}^k), \\ \max\left[1 - \mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)\right] & \text{if } \mu_{A^j}(\bar{\mathbf{x}}) > \mu_{B^j}(\bar{y}^k). \end{cases} \tag{44}$$

It is easy to show that (44) can be replaced by

$$\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k) = \min\left[1, \max\left[1 - \mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)\right] + \rho\left(\mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)\right)\right], \tag{45}$$

where

$$\rho(a,b) = \begin{cases} 1 & \text{if } a \leq b, \\ 0 & \text{if } a > b. \end{cases} \tag{46}$$

Thus, from (17), (39), (36) and (45), the OCFS system based on the Fodor implication is described by

$$\bar{y} = \frac{\sum\limits_{k=1}^{N} \bar{y}^k \mathop{T}\limits_{\substack{j \neq k \\ j=1}}^{N} \min\left[1, \max[1 - \mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}] + \rho(\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k})\right]}{\sum\limits_{k=1}^{N} \mathop{T}\limits_{\substack{j \neq k \\ j=1}}^{N} \min\left[1, \max[1 - \mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}] + \rho(\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k})\right]}. \tag{47}$$

From (17), (39), (36) and the following expression (cf. Table 1):

$$\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k) = \begin{cases} 1 & \text{if } \mu_{A^j}(\bar{\mathbf{x}}) \leq \mu_{B^j}(\bar{y}^k), \\ 0 & \text{if } \mu_{A^j}(\bar{\mathbf{x}}) > \mu_{B^j}(\bar{y}^k), \end{cases} \tag{48}$$

we obtain the description of the OCFS system based on the Sharp implication:

$$\bar{y} = \frac{\sum\limits_{k=1}^{N} \bar{y}^k \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \rho\left(\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}\right)}{\sum\limits_{k=1}^{N} \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \rho\left(\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}\right)}. \tag{49}$$

In much the same way, from (17), (39) and (36), we get the following description of the OCFS system based on the Goguen implication (cf. Table 1):

$$\bar{y} = \frac{\sum\limits_{k=1}^{N} \bar{y}^k \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \min\left[1, \dfrac{p_{j,k}}{\mu_{A^j}(\bar{\mathbf{x}})}\right]}{\sum\limits_{k=1}^{N} \mathop{T}\limits_{\substack{j=1 \\ j \neq k}}^{N} \min\left[1, \dfrac{p_{j,k}}{\mu_{A^j}(\bar{\mathbf{x}})}\right]}. \tag{50}$$

If $\mu_{A^j}(\bar{\mathbf{x}}) = 0$, then $\min[1, p_{j,k}/\mu_{A^j}(\bar{\mathbf{x}})] = 1$ in (50).

For the Gödel fuzzy implication (cf. Table 1), we have the following expression:

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \begin{cases} 1 & \text{if} \quad \mu_{A^j}\left(\bar{\mathbf{x}}\right) \le \mu_{B^j}\left(\bar{y}^k\right), \\ \mu_{B^j}\left(\bar{y}^k\right) & \text{if} \quad \mu_{A^j}\left(\bar{\mathbf{x}}\right) > \mu_{B^j}\left(\bar{y}^k\right), \end{cases} \tag{51}$$

which can be replaced, analogously to the case of the Fodor implication, by

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \min\left[1, \mu_{B^j}(\bar{y}^k) + \rho\big(\mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)\big)\right]. \tag{52}$$

Thus, from (17), (39), (36) and (52), the description of the OCFS system based on the Gödel implication is expressed as follows:

$$\bar{y} = \frac{\sum\limits_{k=1}^{N} \bar{y}^k \underset{\substack{j=1 \\ j\neq k}}{\overset{N}{T}} \min\left[1, p_{j,k} + \rho\big(\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}\big)\right]}{\sum\limits_{k=1}^{N} \underset{\substack{j=1 \\ j\neq k}}{\overset{N}{T}} \min\left[1, p_{j,k} + \rho\big(\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}\big)\right]}. \tag{53}$$

For the OCFS system based on the Yager implication (cf. Table 1), from (17), (39) and (36), we obtain the following equation which describes the system:

$$\bar{y} = \frac{\sum\limits_{k=1}^{N} \bar{y}^k \underset{\substack{j=1 \\ j\neq k}}{\overset{N}{T}} (p_{j,k})^{\mu_{A^j}(\bar{\mathbf{x}})}}{\sum\limits_{k=1}^{N} \underset{\substack{j=1 \\ j\neq k}}{\overset{N}{T}} (p_{j,k})^{\mu_{A^j}(\bar{\mathbf{x}})}}. \tag{54}$$

Now let us consider the OCFS neuro-fuzzy systems based on Zadeh and Wilmott implications. The descriptions of these systems can be determined from (17), (40), (36), and the following expressions (cf. Table 1):

$$\mu_{A^j \to B^j}\left(\bar{\mathbf{x}}, \bar{y}^k\right) = \max\left[\min[\mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)], 1 - \mu_{A^j}(\bar{\mathbf{x}})\right] \tag{55}$$

for the Zadeh implication, and

$$\begin{aligned} \mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k) = \min\Big[&\max\left[1 - \mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)\right], \\ &\max\left[\mu_{A^j}(\bar{\mathbf{x}}), 1 - \mu_{B^j}(\bar{y}^k),\right. \\ &\left.\min\left[1 - \mu_{A^j}(\bar{\mathbf{x}}), \mu_{B^j}(\bar{y}^k)\right]\right]\Big] \end{aligned} \tag{56}$$

for the Wilmott implication. Therefore, the descriptions of these systems are given by (17), where

$$\begin{aligned} \lambda_k\left(\bar{\mathbf{x}}, \bar{y}^k\right) = T\Big\{ &\max\left[\mu_{A^k}(\bar{\mathbf{x}}), 1 - \mu_{A^k}(\bar{\mathbf{x}})\right], \\ &\underset{\substack{j=1 \\ j\neq k}}{\overset{N}{T}} \max\left[\min[\mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}], 1 - \mu_{A^j}(\bar{\mathbf{x}})\right]\Big\} \end{aligned} \tag{57}$$

for systems based on the Zadeh implication, and

$$\lambda_k(\bar{\mathbf{x}}, \bar{y}^k) = T\Big\{ \max\big[\mu_{A^k}(\bar{\mathbf{x}}), 1 - \mu_{A^k}(\bar{\mathbf{x}})\big], \underset{\substack{j=1 \\ j \neq k}}{\overset{N}{T}} \min\Big[\max[1 - \mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}], \max\big[\mu_{A^j}(\bar{\mathbf{x}}), 1 - p_{j,k}, \min[1 - \mu_{A^j}(\bar{\mathbf{x}}), p_{j,k}]\big]\Big]\Big\} \tag{58}$$

for systems based on the Wilmott implication.

The multi-layer connectionst architectures of the systems presented in this section have the form of the architecture illustrated in Fig. 2. The first layer (antecedent fuzzy set layer), the aggregation layer and the last (defuzzification) layer are the same in each architecture. The inference layer is different in each of them. It is easy to construct this layer for each system considered in this section, based on the expressions which depict $\mu_{A^j \to B^j}(\bar{\mathbf{x}}, \bar{y}^k)$ for the corresponding fuzzy implication. Examples of these architectures can be found in (Nowicki, 1999; Nowicki and Rutkowska, 2000b; 2000c; Rutkowska and Nowicki, 2000). The inference layers may contain elements which perform the complement operation (cf. Section 3.2), minimum and/or maximum operation, power operation (for the system based on the Yager implication), the operator defined by (46), and others.

Let us notice that, similarly to the systems under Mamdani's approach, if $p_{j,k} = 0$, for $j, k = 1, \ldots, N$, then the OCFS systems presented here reduce to the NOCF systems under the logical approach described in Section 3.3.

4. Learning Methods

The multi-layer connectionist architectures of the neuro-fuzzy systems presented in Section 3 can be trained similarly to the neural network learning, as mentioned at the beginning of this section. The learning algorithms of special types of the neuro-fuzzy systems described in Section 3 are presented in (Rutkowska, 1997; Rutkowska *et al.*, 1997; Wang, 1994). These algorithms employ the idea of the steepest-descent optimisation technique, analogously to the back-propagation method which is commonly used in order to train artificial neural networks. The FLiNN programme (Piliński, 1997a) is an example of the software which performs training of the neuro-fuzzy systems with the use of this method. It is worth emphasizing that this kind of software does not require mathematical formulae of the recursions which depict the learning algorithm for particular neuro-fuzzy systems. The FLiNN programme realizes the back-propagation learning algorithm for a given neuro-fuzzy multi-layer architecture. This means that this programme can propagate the output error from the last layer of the network to previous layers, and tune the parameters, i.e. the centres and widths of the membership functions, to minimize the output error of the system. It is very important that this programme can realize the learning of the system based on the architecture, i.e. the elements of particular layers and the connections between them.

For some neuro-fuzzy systems it is very difficult to determine a mathematical formula of the learning algorithm. It is much easier to construct a connectionist architecture and to apply the FLiNN programme in order to train the system. For the details concerning the FLiNN software, one can refer to (Piliński, 1997a; 1997b).

The idea of the back-propagation algorithm (Żurada, 1992), employed to tune the parameter σ, is expressed by the recursion

$$\sigma(t+1) = \sigma(t) - \eta \frac{\partial Q(\bar{\mathbf{x}}, d; t)}{\partial \sigma(t)}, \tag{59}$$

where η specifies the speed of learning, Q denotes a measure of the error, $\bar{\mathbf{x}}$ is the input value and d signifies the desired output value. The error is defined by

$$Q(\bar{\mathbf{x}}, d) = \frac{1}{2}\left[\bar{y}(\bar{\mathbf{x}}) - d\right]^2, \tag{60}$$

where $\bar{y}$ is the output value. If we apply this method to the connectionist multi-layer architectures, each elements of these networks propagates the error from the output to the inputs. According to (59), it is necessary to compute the derivative. For the elements of the architectures which implement nondifferentiable functions, such as 'minimum', 'maximum' or 'power', special methods of computing the derivative of nondifferentiable functions are employed. Pairs of the input and desired output values, $(\bar{\mathbf{x}}, d; t)$, $t = 0, 1, 2, \ldots, M$, form a learning sequence. Usually, a similar sequence of data is used to test the system after the training process.

The learning algorithm can be supported by other methods, such as competitive learning, clustering algorithms or genetic algorithms. The more elements the neuro-fuzzy architecture contains, the longer and more difficult the training process is. Therefore, additional methods which support the learning process are very helpful in some applications.

5. Applications

Neuro-fuzzy systems are usually applied to control and classification tasks. In this section, we present results of application of the systems described in Section 3 and trained using methods delineated in Section 4 to this kind of problems.

5.1. Classification

At first, we have chosen a very simple but illustrative example of classification (Nowicki, 1999; Nowicki and Rutkowska, 1999; 2000b; Rutkowska and Nowicki, 2000; Rutkowska *et al.*, 1999). The task is to classify the points located in the square area to three different classes: two semi-rings and the area of the square beyond the semi-rings. The points which belong to the region of the first semi-ring should be assigned number 1. The points included in the second semi-ring ought to be assigned to value -1. The points located in the square area but beyond the regions of both the semi-rings should be associated with number 0. In order to perform the classification task, a learning sequence of 1089 points, evenly placed on the square area, with properly associated numbers 1, -1 or 0, was created. The neuro-fuzzy systems applied ten rules in the form (1). Five of them referred to the first semi-ring class (the centres of the consequent fuzzy sets equaled 1) and five to the second semi-ring class (the centres of the consequent fuzzy sets equaled -1). The two coordinate values of the points in the square area were fed to the inputs of the systems. The output represents the number of the class which corresponds to a classified point. After the learning process, the systems were tested on the sequence of 4225 points, evenly placed in the square area. The results of the classification, performed by different neuro-fuzzy systems presented in this paper, are reported in Fig. 11. As we see, the worst effect of the classification was obtained for the systems which employ Mamdani's approach to fuzzy inference, i.e. NOCFS and OCFS systems based on Mamdani and Larsen's implications (see Figs. 11(a), (e) and (f)). The NOCFS systems which applied Mamdani and Larsen's relations correctly classified the points located in the area of the two semi-rings but the points placed beyond these regions were classified to the nearest semi-ring. This is illustrated in Fig. 11(a). This system requires more rules, and some of them should refer to the third class (the area beyond the semi-rings) in order to perform a good classification. The NOCFS neuro-fuzzy systems based on the logical approach to fuzzy inference perform much better in the case of ten rules associated with two semi-ring classes. These systems can recognize the third class. This result is illustrated in Fig. 11(b), (c) and (d). The OCFS systems employed the rules with different consequent fuzzy sets. The centres of the fuzzy sets were trained, so these rules did not refer strictly to the semi-ring classes. Figures 11(g)–(o) present better results than Figs. 11(e) and (f), which means that the systems based on the logical approach perform better in this application than those based on Mamdani's approach.

We have also applied neuro-fuzzy systems to the well-known Iris classification task. There are three species of the iris flowers: *Setosa*, *Versicolor* and *Virginica.* Fisher's iris data (Fisher, 1936) were used. Each of the data vectors was composed of four components which corresponded to four features of the iris flower: *sepal length*, *sepal width*, *petal length* and *petal width.* There were 150 data vectors, 50 for each of the iris species. The task was to classify the data to the proper class of the three iris species. The NOCFS neuro-fuzzy systems were applied to solve the classification problem. Table 2 presents results of the classification for the systems constructed based on two and three fuzzy IF-THEN rules. The table shows that the results are better for the systems which employ the logical approach to fuzzy inference. The first row refers to the NOCFS neuro-fuzzy systems based on Mamdani and Larsen's fuzzy implications (relations), i.e. the systems with Mamdani's approach. The percentage of the correct classifications is lower than in the case of the logical approach, i.e. the systems based on the Kleene-Dienes, Łukasiewicz, Reichenbach, Fodor implications (the second row in the table), the systems based on the Sharp, Goguen, Gödel, Yager implications (the third row), and the systems based on Zadeh and Wilmott's implications (the last row).

The next examples of classification concern medical diagnosis. We have applied neuro-fuzzy systems to the problems of diagnosing a tumor of the mucous membrane of uterus and breast cancer. The data for the former case were received from a hospital

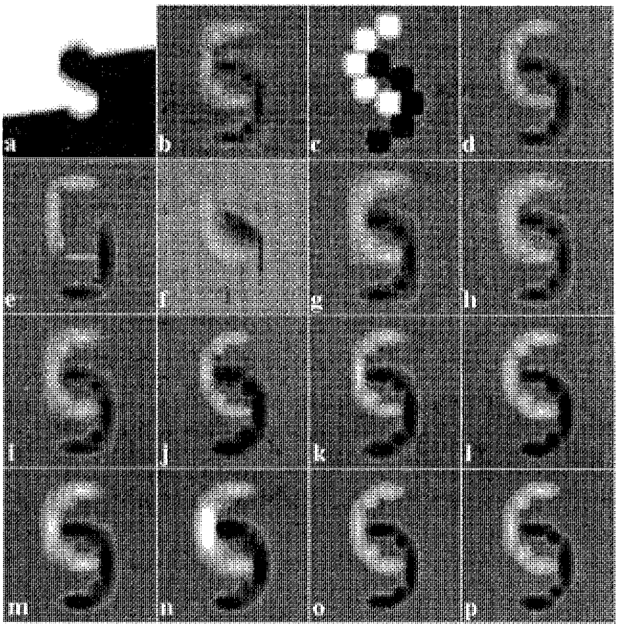

Fig. 11. Results of the semi-ring classification problem obtained using NOCFS systems based on: (a) Mamdani, Larsen, (b) Kleene-Dienes, Łukasiewicz, Reichenbach, Fodor, (c) Sharp, Goguen, Gödel, Yager, (d) Zadeh, Wilmott implications, and OCFS systems based on: (e) Larsen, (f) Mamdani, (g) Kleene-Dienes, (h) Łukasiewicz, (i) Reichenbach, (j) Fodor, (k) Sharp, (l) Goguen, (m) Gödel, (n) Yager, (o) Zadeh, (p) Wilmott implications.

Table 2. Results of the iris classification.

NOCFS systems	for 2 rules	for 3 rules
Mamdani, Larsen	66.67%	97.33%
Kleene-Dienes, Łukasiewicz, Reichenbach, Fodor	97.33%	98.00%
Sharp, Goguen, Gödel, Yager	67.33%	98.00%
Zadeh, Wilmott	97.33%	98.00%

in Częstochowa, Poland (Rutkowska, 1997; 1998). The training sequence contained 54 and the testing sequence 11 data records of women. Each record collected 9 attributes: the period of time after the menopause, body mass index, luteinizing hormone, follicle-stimulating hormone, prolactin, estrone, estradiol, aromatase and estrogenic receptor. The data for the latter case were provided by the University of Wisconsin Hospitals and availaible from the Internet (Mertez and Murphy, 2000; Wolberg and Mangasarian, 1990). From this database we got 487 different values of 10 attributes: clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, mitoses and the diagnosis. The data were split into learning and testing sequences of 387 and 100 records, respectively. In both the cases the attributes were treated as inputs to the neuro-fuzzy systems, and the diagnosis as the output. In the former example we obtained 100% correct answers inferred by all the neuro-fuzzy systems (Nowicki and Rutkowska, 2000a; 2000c; Rutkowska and Nowicki, 2000), whereas in the latter example the percentage of the correct system responses ranged from 97,87% to 98,72%, depending on the kind of the neuro-fuzzy system employed.

5.2. Control

Fuzzy and neuro-fuzzy systems constructed using Mamdani's approach have been mostly applied as fuzzy and neuro-fuzzy controllers (see e.g. Wang, 1994). Therefore, we have checked the performance of the neuro-fuzzy systems, presented in this paper, when applied to control problems. Well-known examples of control of the truck backer-upper and the inverted pendulum were choosen (see e.g. Rutkowska, 1997; Rutkowska *et al.*, 1997; Wang, 1994). Figure 12 illustrates the trajectories of the truck controlled by different neuro-fuzzy systems.

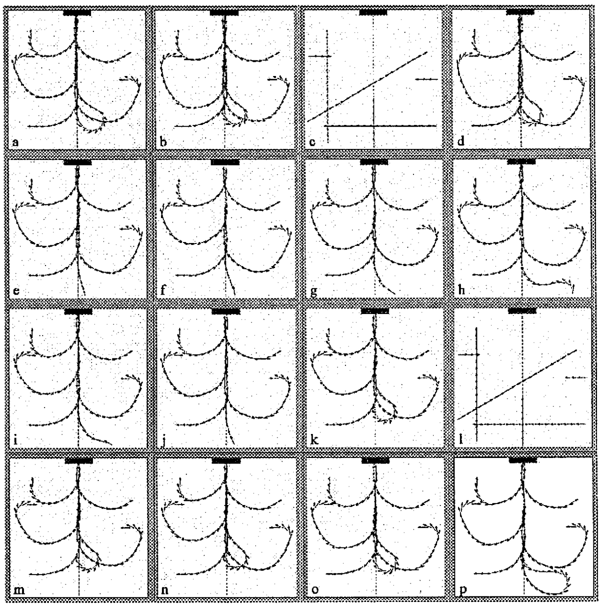

Fig. 12. Results for the truck backer-upper control problem obtained using NOCFS systems based on: (a) Mamdani, Larsen, (b) Kleene-Dienes, Łukasiewicz, Reichenbach, Fodor, (c) Sharp, Goguen, Gödel, Yager, (d) Zadeh, Wilmott implications, and OCFS systems based on: (e) Larsen, (f) Mamdani, (g) Kleene-Dienes, (h) Łukasiewicz, (i) Reichenbach, (j) Fodor, (k) Sharp, (l) Goguen, (m) Gödel, (n) Yager, (o) Zadeh, (p) Wilmott implications.

As can be seen from Fig. 12, only the systems based on the Goguen fuzzy implication cannot successfully control the truck. Some of the OCFS systems work worse than their NOCF counterparts. This is not because of limited capabilities of these systems. The OCFS neuro-fuzzy architectures contain more elements, so it is more difficult to tune their parameters during the learning process.

6. Conclusions

The paper presents neuro-fuzzy systems built as multi-layer connectionist networks. This kind of architectures can be trained to tune the parameters of the membership functions using the software, such as the FLiNN programme (Piliński, 1997a). Consequently, the learning algorithm which incorporates the idea of back-propagation is realised for each neuro-fuzzy system using its architecture. The main advantage is that this method does not require mathematical forms of the recursions which describe the algorithm of tuning system parameters. Thus, we can change the architectures and the programme can easily perform the learning algorithm. Even if the architectures contain many elements, they can be trained in this way. Let us notice that in many cases it is difficult to determine the mathematical recursions of the back-propagation learning, especially for nondifferentiable functions.

The most popular and commonly used Mamdani's approach to fuzzy inference is compared in this paper with the logical approach. The neuro-fuzzy systems constructed using both the approaches have been applied to classification and control problems. It seems that the Mamdani's approach is more suitable for control tasks, while the logical approach performs better in classification tasks. However, both the approaches can be employed in both the applications.

Looking at the results of the applications of the systems based on different fuzzy implications, we observe a difference in their performance. More detailed information about the performance of various neuro-fuzzy systems can be found in (Rutkowska *et al.*, 2000). Let us notice that the systems based on some implications do not perform well for Gaussian membership functions. In these cases triangular membership functions should be applied.

The NOCFS and OCFS systems have been considered in this paper. The former ones are a special case of the latter ones. Thus, those of the NOCFS neuro-fuzzy systems are simpler than the architectures of the OCFS systems. The most commonly applied neuro-fuzzy systems based on the Mamdani or the Larsen's fuzzy relation (Rutkowska, 1997; Rutkowska *et al.*, 1997; Wang, 1994) are NOCFS systems, with the Mamdani's approach.

In (Czogała and Łęski, 2001) a specific type of the equivalence of inference results using a special type of fuzzy implication and the Mamdani/Larsen relation is studied. The authors conclud that the inference algorithms based on conjuctive operators (Mamdani, Larsen) in some cases seem to be faster, simpler and more exact than the inference based on fuzzy implications (logical approach), but the latter type of inference is sounder from the logical point of view.

References

Cordon O., Herrera A. and Peregrin A. (1997a): *Applicability of the fuzzy operators in the design of fuzzy logic controllers.* — Fuzzy Sets Syst., Vol.86, No.1, pp.15–41.

Cordon O., Herrera A. and Peregrin A. (1997b): *Searching for Basic Properties Obtaining Robust Implication Operators in Fuzzy Control.* — Techn. Rep. Dept. Comp. Sci. Artif. Intell.; available at `http://decsai.ugr.es/~ocordon/public.html`.

Czogała E., Fodor J. and Łęski J. (1997a): *On modeling destructive logical type systems using Fodor fuzzy implication.* — Proc. 5th Europ. Congress *Intelligent Techniques and Soft Computing (EUFIT'97)*, Aachen, Germany, pp.340–343.

Czogała E. and Łęski J. (1997): *On destructive fuzzy logic controllers.* — Proc. 5th Zittau Fuzzy Colloquium, Zittau, Germany, pp.8–12.

Czogała E. and Łęski J. (2000): *Fuzzy and Neuro-Fuzzy Intelligent Systems.* — Heidelberg: Physica-Verlag.

Czogała E. and Łęski J. (2001): *On equivalence of approximate reasoning results using different interpretations of fuzzy if-then rules.* — Fuzzy Sets Syst., Vol.117, No.2, pp.279–296.

Driankov D., Hellendoorn H. and Reinfrank M. (1993): *An Introduction to Fuzzy Control.* — Berlin: Springer-Verlag.

Filev D.P. and Yager R.R. (1995): *Simplified methods of reasoning in fuzzy models.* — Proc. 1995 IEEE Int. Conf. *Fuzzy Systems*, Yokohama, Japan, pp.123–129.

Fisher R.A. (1936): *The use of multiple measurements in taxonomic problems.* — Ann. Eugenics, Vol.7, pp.179–188.

Lee C.C. (1990): *Fuzzy logic in control systems: fuzzy logic controllers - part II.* — IEEE Trans. Syst. Man Cybern., Vol.20, No.2, pp.419–435.

Mertez C.J. and Murphy P.M. (2000): *UCI respiratory of machine learning databases.* — `http://www.ics.uci.edu/pub/machine-learning-databases`.

Moody J. and Darken C. (1989): *Fast learning in networks of locally-tuned processing units.* — Neural Computn., Vol.1, No.2, pp.281–294.

Nowicki R. (1999): *Neuro-Fuzzy Systems Performing Different Types of Fuzzy Inference.* — Ph.D. Thesis, Technical University of Częstochowa, Częstochowa, Poland (in Polish).

Nowicki R. and Rutkowska D. (1999): *Fuzzy inference neural networks based on destructive and constructive approaches and their application to classifications.* — Proc. 4th Nat. Conf. *Neural Networks and Their Applications*, Zakopane, Poland, pp.302–307.

Nowicki R. and Rutkowska D. (2000a): *New neuro-fuzzy architectures.* — Proc. Int. Conf. *Artificial and Computational Intelligence for Decision, Control and Automation in Engineering and Industrial Applications (ACIDCA'2000)*, Intelligent Methods, Monastir, Tunisia, pp.82–87.

Nowicki R. and Rutkowska D. (2000b): *Neuro-fuzzy architectures based on Yager implication.* — Proc. 5th Nat. Conf. *Neural Networks and Soft Computing*, Zakopane, Poland, pp.353–360.

Nowicki R. and Rutkowska D. (2000c): *Neuro-fuzzy systems based on Gödel and Sharp implication.* — Proc. Int. Conf. *Application of Fuzzy Systems and Soft Computing (ICAFS-2000)*, Siegen, Germany, pp.232–237.

Piliński M. (1997a): *FLiNN - User Manual.* — Częstochowa: Polish Neural Network Society (in Polish).

Piliński M. (1997b): *Methodology of Authomatic Learning of Fuzzy Controllers Using Neural Networks.* — Ph.D. Thesis, Technical University of Częstochowa, Częstochowa, Poland (in Polish).

Rutkowska D. (1997): *Intelligent Computational Systems.* — Warsaw: PLJ Academic Publishing House (in Polish).

Rutkowska D. (1998): *Fuzzy neural networks with an application to medical diagnosis.* — Biocybern. Biomed. Eng., Vol.18, No.1–2, pp.71–78.

Rutkowska D. and Nowicki R. (1999): *Constructive and destructive approach to neuro-fuzzy systems.* — Proc. EUROFUSE-SIC'99, Budapest, Hungary, pp.100–105.

Rutkowska D. and Nowicki R. (2000): *Neuro-fuzzy architectures based on Fodor implication.* — Proc. 8th Zittau Fuzzy Colloquium, Zittau, Germany, pp.230–237.

Rutkowska D., Piliński M. and Rutkowski L. (1997): *Neural Networks, Genetic Algorithms, and Fuzzy Systems.* — Warsaw: PWN (in Polish).

Rutkowska D., Nowicki R. and Rutkowski L. (1999): *Neuro-fuzzy system with inference process based on Zadeh implication.* — Proc. 3rd Int. Conf. *Parallel Processing & Applied Mathematics (PPAM'99)*, Kazimierz Dolny, Poland, pp.597–602.

Rutkowska D., Nowicki R. and Rutkowski L. (2000): *Neuro-fuzzy architectures with various implication operators.* — Proc. Int. Symp. *Computational Intelligence (ISCI 2000)*, Košice, Slovakia, pp.214–219.

Wang L.-X. (1994): *Adaptive Fuzzy Systems and Control.* — Englewood Cliffs, NJ: PTR Prentice Hall.

Wolberg W.H. and Mangasarian O.L. (1990): *Multisurface method of pattern separation for medical diagnosis applied to breast cytology.* — Proc. Nat. Acad. Sci., USA, Vol.87, pp.9193–9196.

Yager R.R. and Filev D.P. (1994): *Essential of Fuzzy Modeling and Control.* — New York: Wiley.

Zadeh L.A. (1973): *Outline of a new approach to the analysis of complex systems and decision processes.* — IEEE Trans. Syst. Man Cybern., Vol.3, No.1, pp.28–44.

Żurada J.M. (1992): *Introduction to Artificial Neural Systems.* — New York: West Publishing Company.