WROCLAW UNIVERSITY OF TECHNOLOGY

# Development of a clinical decision support system, using rough sets and type 2 fuzzy control

by

Joel Quesada Vallejo

in the

Faculty of Computer Science and Management

Institute of Informatics

June 2015

WROCLAW UNIVERSITY OF TECHNOLOGY

# *Abstract*

Faculty of Computer Science and Management
Institute of Informatics

Computer engineering

by Joel Quesada Vallejo

In this thesis, the specification of a histopathology decision making support system, based on Pawlak's information system concept in combination with 2-Type fuzzy reasoning is presented. The proposed system supports the recognition process of HER-2/neu histopathology preparations through microscopy image information analysis. Pawlak's information system is used to identify the decisive set of features and the optimal set of decision rules under the considered histopathology problem. Next, the so generated decision rules will be transformed into fuzzy rules and exploited using 2-Type fuzzy reasoning.

WROCLAW UNIVERSITY OF TECHNOLOGY

# *Abstract*

Faculty of Computer Science and Management
Institute of Informatics

Computer engineering

by Joel Quesada Vallejo

En esta tesis, la especificación de un sistema de soporte a la toma de decisiones, basado en el concepto del sistema de información de Pawlak con combinación de razonamiento fuzzy de Tipo-2 es presentado. El sistema propuesto colabora con el proceso de identificación de HER-2/neu a través del análisis de imágenes microscópicas de membranas celurares. El sistema de información de Pawlak es utilizado para identificar el conjunto decisivo de características y el conjunto óptimo de reglas para el considerado problema. Finalmente, el conjunto de reglas generado es transformado en reglas fuzzy y explotado por el mecanismo de inferencia basado en razonamiento fuzzy de Tipo-2.

WROCLAW UNIVERSITY OF TECHNOLOGY

# *Abstract*

Faculty of Computer Science and Management
Institute of Informatics

Computer engineering

by Joel Quesada Vallejo

En aquesta tesi, l'especificació d'un sistema de suport a la presa de decisions, basat en el concepte del sistema d'informació de Pawlak amb combinació de raonament fuzzy de Tipus-2 és presentat. El sistema proposat col·labora amb el procés d'identificació d'HER2/neu a través de l'anàl·lisis d'imatges microscòpiques de membranes cel·lulars. El sistema d'informació de Pawlak és utilitzat per identificar el conjunt decisiu de caracaterístiques i el conjunt òptim de regles pel considerat problema. Finalment, el conjunt de regles generat és transformat en regles fuzzy i explotat pel mecanisme d'inferència basat en raonament fuzzy de Tipus-2.

# Contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| **HER2** | **H**uman **E**pidermal **G**rowth factor 2 |
| **FISH** | **F**luorescence **I**n **S**itu **H**ybridisation |
| **RGB** | **R**ed **G**reen **B**lue colour model |
| **HSV** | **H**ue **S**aturation **V**alue colour model |
| **T1 FS** | **T**ype-1 **F**uzzy **S**et |
| **IT2 FS** | **I**nterval **T**ype-2 **F**uzzy **S**et |
| **AI** | **A**rtificial **I**ntelligence |
| **FLS** | **F**uzzy **L**ogic **S**ytem |

# Symbols

| | | |
|---|---|---|
| $\sigma$ | standard deviation | |
| $x_0$ | expected value | |
| $e$ | Euler's number | 2.71828 |

# Chapter 1

# Introduction

A **clinical decision-support system** is any computer program designed to help health-care professionals to make clinical decisions. In a sense, any computer system that deals with clinical data or knowledge is intended to provide decision support.[1]

In this thesis a clinical decision support system, designed to assist physicians and other health professionals with decision making tasks of establishing a diagnosis on the basis of patient histopathology images, is presented.

## 1.1 Practical problem involved

Breast cancer is considered as the most common cancer in women and is the first cause of morbidity due to cancer in women age 20 to 59 years. In approximately 20% of the diagnosed breast cancers, an overexpression of Human Epidermal Growth Factor Receptor 2 (HER2/neu) is noted.

### 1.1.1 HER2 positive breast cancer

HER2-positive breast cancer is a breast cancer that tests positive for a protein called human epidermal growth factor receptor 2 (HER2), which promotes the growth of cancer cells.

In about 1 of every 5 breast cancers, the cancer cells have a gene mutation that makes an excess of the HER2 protein. HER2-positive breast cancers tend to be more aggressive

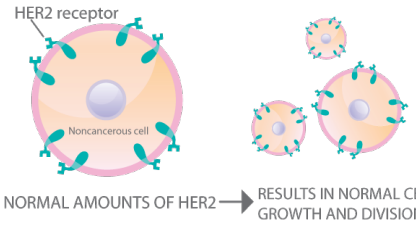



FIGURE 1.1: Normal expression of HER2

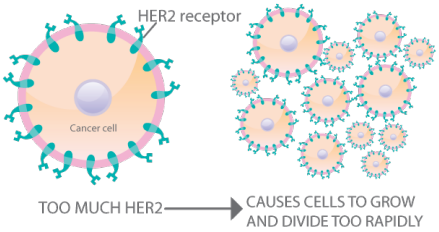



FIGURE 1.2: Overexpression of HER2

than other types of breast cancer. They are less likely to be sensitive to hormone therapy, though many people with HER2-positive breast cancer can still benefit from hormone therapy.[2]

Treatments that specifically target HER2 are very effective. These treatments are so effective that the prognosis for HER2-positive breast cancer is actually quite good. Main example for this kind of treatments is trastuzumab.

### 1.1.2 Diagnostic procedure for HER2 cancer

In the last years a routine diagnostic procedure based on immunohistochemical (IHC) and fluorescent in situ hybridization techniques (FISH) was established to identify potential responders to trastuzumab therapy, which in numerous studies was shown to reduce the risk of recurrence and mortality rate in early and advanced stage breast cancer.

In routine clinical practice, a two-step diagnostic procedure is used for evaluation of HER2 expression status. At first, in IHC HER2 preparations of paraffin-embedded breast cancer specimens a visual semi-quantitative examination of membranous cell staining in tumor cells under a light microscope is performed. For this purpose a categorical classification system based on staining intensity and its pattern is utilized and briefly encoded as shown in Table 1.1.

| Encoding | Meaning |
|---|---|
| 0 | no staining |
| 1+ | weak membrane staining |
| 2+ | nonuniform complete membrane staining |
| 3+ | intense membrane staining |

TABLE 1.1: Categorical classification system based on staining intensity

For trastuzumab treatment cases scored 3+ in the IHC preparations are classified for the therapy, whereas cases scored 2+ are subjected to further testing with costly FISH examination to finally determine HER2 expression status. Therefore, it is **extremely needed** to introduce less expensive diagnostic process for correct recognition of the corresponding HER-2/neu classes.

This thesis tries to give a solution to such a problem with the development of a clinical support system tested over over real clinical data of HER-2/neu breast cancer histopathology images.

## 1.2 System function

Decision-support programs generally fall into two categories: those that assist healthcare workers with determining what is true about a patient (usually what the correct diagnosis is) and those that assist with decisions about what to do for the patient (usually what test to order, whether to treat).

This research is focused in helping the clinicians about what to do for the patient. In our case, to decide whether a patient is appropriate for trastuzumab treatment or not.

We would like to introduce the specification of a histopathology decision making support system, based on Pawlak's information system concept in combination with Type2-fuzzy reasoning.

Given breast cancer histopathology images a image feature extractor is used to extract important features from them and to build a dataset. This dataset has crisp values (features as red, green and blue from the image given). Therefore, as we want these values to be fuzzy, a generalisation is used to transform these values into fuzzy values.

Over such a decision table, Pawlak's data mining concept is applied (which involves rough sets) in order to generate the optimal set of decision rules.

Next, the so generated decision rules will be transformed into fuzzy rules and exploited in fuzzy reasoning.

The proposed system is planned to support the recognition process of HER-2/neu histopathology preparations through microscopy image information analysis.

# Chapter 2

# Problem and formulation of the concept

This research tries to give a solution to a classical Data Mining problem. Data Mining is defined as the procedure of extracting information from huge sets of data. In other words, we can say that data mining is mining knowledge from data. There are two forms of data analysis that can be used for extracting models describing important classes or to predict future data trends. These two forms are as follows: classification and prediction. Classification models predict categorical class labels; and prediction models predict continuous valued functions. For example, we can build a classification model to categorize bank loan applications as either safe or risky, or a prediction model to predict the expenditures in dollars of potential customers on computer equipment given their income and occupation.

In our case, we are trying to extract knowledge from clinical data of HER-2/neu breast cancer histopathology images. This knowledge is used by an inference mechanism to given a new HER-2/neu breast cancer histopathology image decide the treatment that should be used in a patient. Therefore, we are dealing with a classification problem.

## 2.1 Proposition to solve the problem

This research is focused on create an Intelligent Support System using rough sets and 2-Type Fuzzy sets. An Image Feature Extractor is needed to get the most important

features from HER-2/neu breast cancer histopathology images. Therefore, we can divide the implementation of our system in four different interfaces:

1. **Image feature extractor.** Extracts the corresponding image features from HER-2/neu breast cancer histopathology images.
2. **Fuzzyfication Interface.** Fuzzifies our image features. Generates the corresponding decision table assuming:
   - Cells as objects.
   - Image features as attributes.
   - Attribute values as fuzzy sets (including the decision attribute).
3. **Pawlak's Information System.** Eliminates object conflicts, generate attributes reduct and generates the optimal set of decision rules.
4. **Inference mechanism.** Define common output of the system interpretable as a decision making support.

Following subsections explain in more detail these interfaces. However, we will not get so much into detail with the theory involved. Deep explanation of the theory involved on it will be followed in next chapters.

### 2.1.1 Image feature extractor

In the presented system, the input is a set of histopathology images. Each hispathology image contains different cells. These cells are classified in two different classes regarding their aggressiveness: low aggressiveness and high aggressiveness. All the information regarding to the cells (coordinates and relation between axis) is provided. This interface is used to build a dataset that contains cells as a objects and image feature as attributes of these cells.

### 2.1.2 Fuzzyfication interface

This interface is designed to convert controller inputs into information that the inference mechanism can easily use to activate and apply rules.

The input of this interface is the dataset given by the Image feature extractor mentioned above. This dataset has crisp values. Instead, we would like to have linguistic variables, concretely we use three different linguistic values: low, normal and high to describe all our features.

We use the following fuzzification concept: for every image feature over the considered learning set, a Gaussian distribution was generated, which was used to define the fuzzy set 'medium' over each image feature:

$$\mu_{medium} = e^{\frac{-(x-x_0)^2}{2\sigma^2}} \pm offset \tag{2.1}$$

Next, the fuzzy sets 'low' and 'high' were defined as follows:

$$\mu_{low} = \begin{cases} 1 - e^{\frac{-(x-x_0)^2}{2\sigma^2}} \pm offset & : \mathrm{x} < \mathrm{x}_0 \\ 0 \pm offset & : \mathrm{x} \geq \mathrm{x}_0 \end{cases} \tag{2.2}$$

$$\mu_{high} = \begin{cases} 0 \pm offset & : \mathrm{x} \leq \mathrm{x}_0 \\ 1 - e^{\frac{-(x-x_0)^2}{2\sigma^2}} \pm offset & : \mathrm{x} > \mathrm{x}_0 \end{cases} \tag{2.3}$$

Notice that the membership function give us a interval. This means that we are using 2-Type Fuzzy Sets.

### 2.1.3 Pawlak's Information System

Pawlak's Information Systems Theory, based on rough sets concepts is applied in this interface.

This interface has three main functions:

- Eliminate object conflicts, using some kind of approximation - we will talk about it in next chapters.
- Generate the attributes reduct. Special algorithm for this will be used.
- Generate the minimal set of decision rules that correctly cover the decision problem. This means excluding the undifferentiating attributes and/or attribute values and combining information with respect to the decision attribute.

The rules generated by this interface will be used by our Inference Mechanism in order to classify our new data (test dataset). Deeper explanation of the methods used for it are written in following chapters.

### 2.1.4 Inference Mechanism for 2-Type Fuzzy Sets

Finally, the knowledge achieved from our Learning Dataset will be used by the Inference Mechanism. When we talk about knowledge we are refering to the rules that has been generated by our Pawlak's Information System. Fuzzy control of type 2 are used in our interface. Therefore, special algorithms for this kind of fuzzy sets are performed. They will be explained in more detail in following chapters.

## 2.2 Conceptual scheme

This section provides a conceptual scheme to help the reader understand how works our system.

The scheme is divided in two: one for the process performed over the Learning Dataset to get our knowledge database, and another one for the process performed over the Test/Validation Dataset. The image feature extractor interface is skipped in this place. We assume that we already have a dataset splitted in two datasets: Learning and Test Dataset.

Figure 2.1 shows the process performed in our Learning Dataset. The process is performed in two different interfaces: Fuzzyfication Interface and Pawlak's Information System Interface. As an output we get the optimal set of rules regarding our data.

Next, the so-generated rules are used by our Inference Mechanism in order to predict the HER2 overexpression of new cells. As shown in 2.2 the input in this case is a validation (test) dataset. Through a process that will be explained in more detail in next chapters our system is able to give a crisp output (a real number). This output is interpretable as a decision making support.

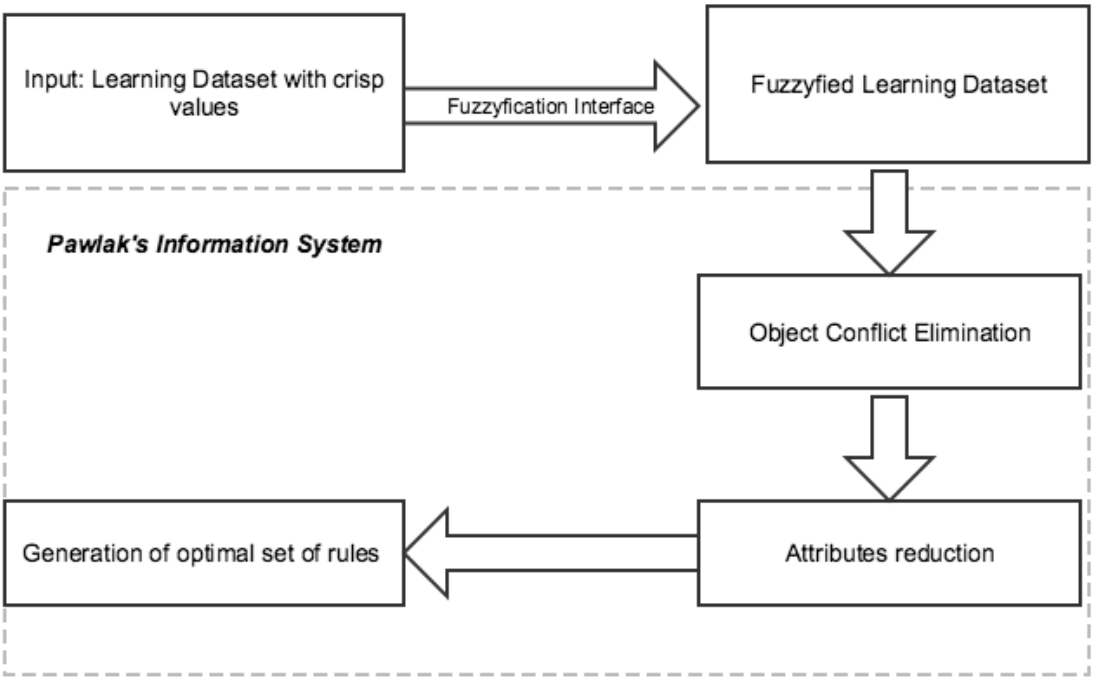



FIGURE 2.1: Rule extraction process

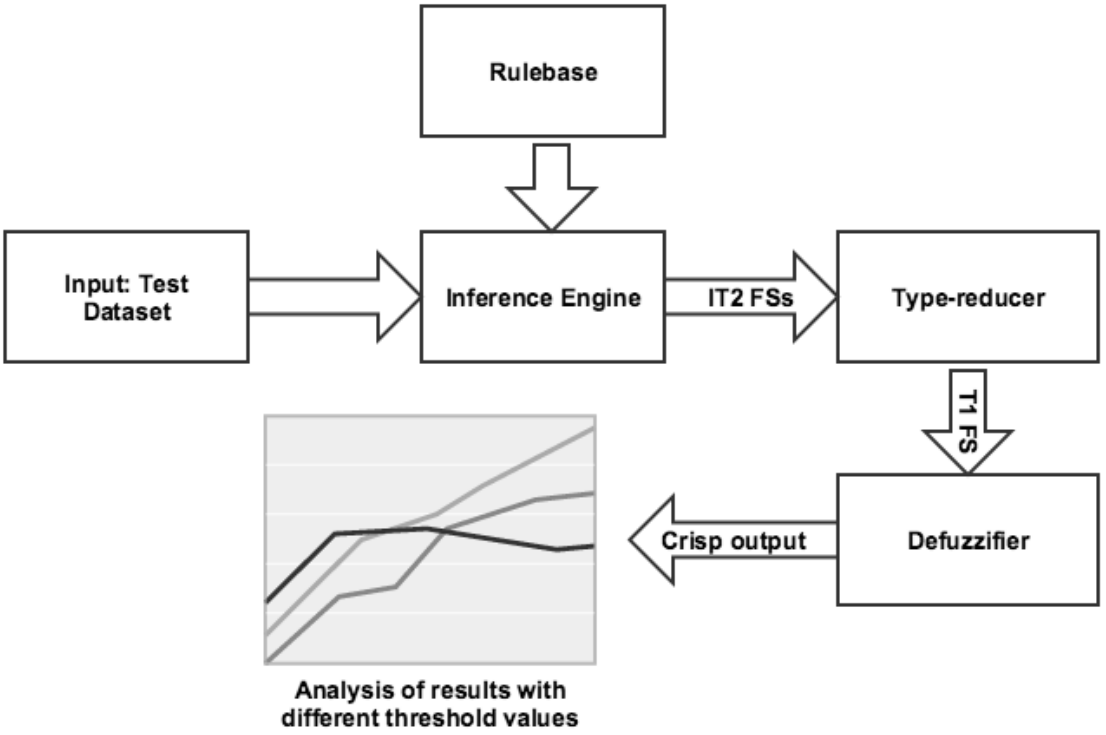



FIGURE 2.2: Decision making process

# Chapter 3

# Theoretical background

In this chapter, the preliminaries of fuzzy sets, fuzzy control of Type-2 and rough sets are described.

## 3.1 Fuzzy sets

A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one. The notions of inclusion, union, intersection, complement, relation, convexity, etc., are extended to such sets, and various properties of these notions in the context of fuzzy sets are established. In particular, a separation theorem for convex fuzzy sets is proved without requiring that the fuzzy sets be disjoint.[5]

### 3.1.1 Type-1 Fuzzy sets

Type-1 fuzzy set (T1 FS) theory was first introduced by Zadeh in 1965 and has been successfully applied in many areas including modeling and control[7], data mining, etc.

#### 3.1.1.1 Formulation of fuzzy sets

Let $X = \{x_1, x_2, ..., x_n\} \subseteq \mathbb{R}$ be some finite set of elements (domain), then we shall call 'A' the fuzzy subset of X, if and only if: $A = \{(x, \mu_A(x)) | x \in X\}$, where $\mu_A$ is a function

that maps X onto the real unit interval [0,1], i.e. $\mu_A : X \to [0,1]$. The function $\mu_A$ is also known as the *membership function* of the fuzzy set A, as its values represents the grade of membership of the elements of X to the fuzzy set A. Here the idea is that we can use membership functions, as characteristic functions (any crisp set can be defined by its characteristic function) for fuzzy, imprecisely described sets. Let A and B be two fuzzy subsets of X, then the basic set operations: *union* and *intersection* of A and B, are defined as follows: $\mu_{A\cup B}(x) = max\{\mu_A(x), \mu_B(x)\}$, $\mu_{A\cap B}(x) = min\{\mu_A(x), \mu_B(x)\}$.

An example of a T1 FS, X, is shown in Figure 3.1. When only integer numbers are considered in the x domain, the T1 FS can be represented as {0/2, 0.5/3, 1/4, 1/5, 0.67/6, 0.33/7, 0/8}, where 0/2 means that number 2 has a *membership degree* of 0 in the T1 FS X, 0.5/3 means number 3 has a membership degree of 0.5 in the T1 FS X, etc.

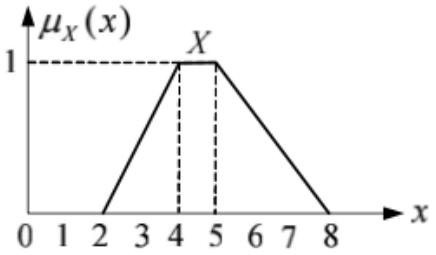



FIGURE 3.1: T1 FS

#### 3.1.1.2 Fuzzy logic systems

A FLS maps crisp inputs into crisp outputs. It contains four components:

- **Rules**. *Rules* may be provided by experts (in our case doctors) or can be extracted from numerical data. In either case, rules are expressed as a collection of **IF-THEN** statements, e.g., **IF** *red* is high and *green* is low **THEN** cancer is aggressive.
- **Fuzzifier**. The *fuzzifier* maps crisp numbers into fuzzy sets. It is needed in order to activate rules which are in terms of linguistic variables, which have fuzzy sets associated with them.
- **Inference Engine**. The *inference* engine of the FLS maps fuzzy sets into fuzzy sets. It handles the way in which rules are combined.

- **Defuzzifier**. The defuzzifer maps output sets into crisp numbers. In our case for example, high values can refer to high cancer HER overexpression.

A Type-1 FLS is depicted in 3.2.[8]

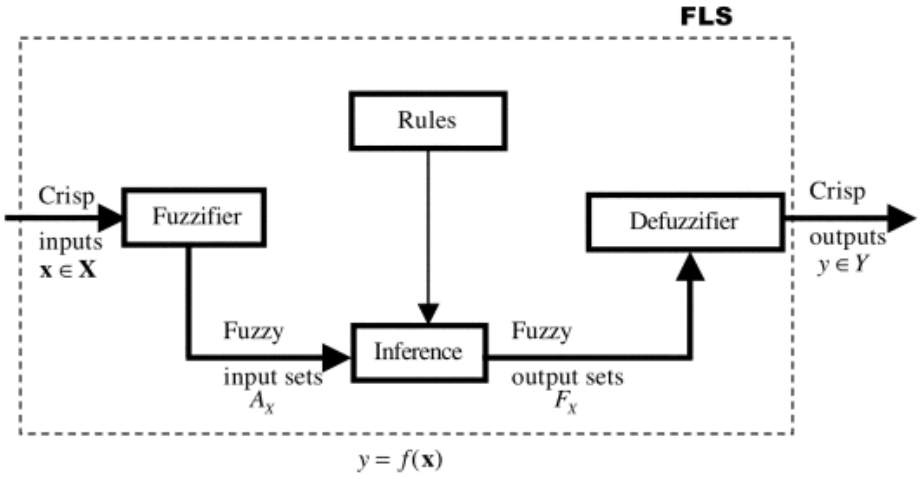



FIGURE 3.2: FLS scheme

### 3.1.2 Type-2 Fuzzy Sets

From the very beginning of fuzzy sets, criticism was made about the fact that the membership function of a type-1 fuzzy set has no uncertainty associated with it, something that seems to contradict the word fuzzy, since that word has the connotation of lots of uncertainty.[4]

This thesis is especially focused in Interval Type-2 Fuzzy Sets. Therefore, explicitly formulations of general Type-2 Fuzzy sets will be avoided. For the reader that is interested in learning about general Type-2 Fuzzy Sets following references are useful. [9–11]

#### 3.1.2.1 Interval Type-2 Fuzzy Sets

Computations using general T2 FSs are very costly. Interval type-2 (IT2) FS , a special case of type-2 FS, are currently the most widely used for their reduced computational cost.

An example of an IT2 FS, $\widetilde{X}$, is shown in 3.3. Observe that unlike a T1 FS, whose membership for each x is a number, the membership of an IT2 FS is an interval. For

example, the membership of number 3 is [0.25, 1], and the membership of number 5 is [0.75, 1].

Observe also that an IT2 FS is bounded from the above and below by IT2 FSs are particularly useful when it is difficult to determine the exact MF, or in modeling the diverse opinions from different individuals. The MFs can be constructed from surveys, or using optimization algorithms, which are called *upper* MF (UMF) and *lower* MF (LMF), respectively. The area between $\overline{X}$ and $\underline{X}$ is the *footprint of uncertainty* (FOU).[3]

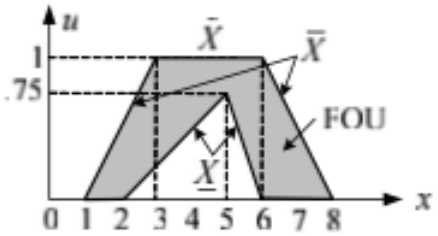



FIGURE 3.3: IT2 FS

### 3.1.2.2 Interval Type-2 Fuzzy Logic Systems

A general T2 FLS is depicted in 3.4. It is very similar to the T1 FLS in 3.2, the major structural difference being that the defuzzifier block of a T1 FLS is replaced by the output processing block in a T2 FLS.[6]

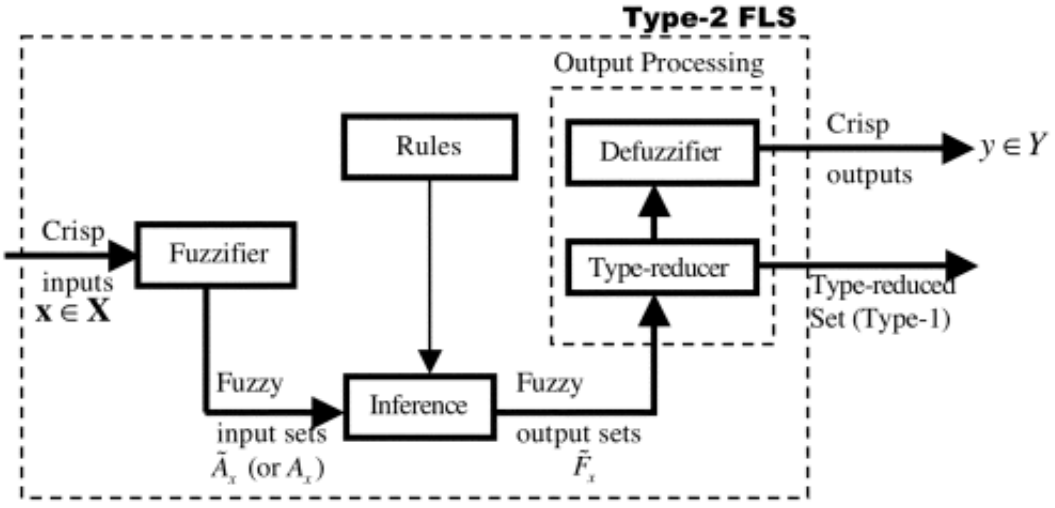



FIGURE 3.4: IT2 FS

In practice the computations in an IT2 FLS can be significantly simplified. Consider the rulebase of an IT2 FLS consisting of $N$ rules assuming the following form:

$R^n$: IF $x_i$ is $\widetilde{X}_1^n$ and $\cdots$ and $x_I$ is $\widetilde{X}_I^n$, THEN $y$ is $Y^n$ $\qquad n$ = 1,2,...,$N$

where $\widetilde{X}_i^n$ (i = 1, ..., $I$) are IT2 FSs, and $Y^n$=[$\underline{y}^n$, $\overline{y}^n$] is an interval.

Assume the input vector is **x'** = ($x'_1$, $x'_2$, ..., $x'_I$ ). Typical computations in an IT2 FLS involve the following steps:

1. Compute the membership of $x'_i$ on each $X_n^i$, $[\mu_{\underline{X}_i^n}(x'_i), \mu_{\overline{X}_i^n}(x'_i)]$, $\quad i = 1, 2, \ldots, I$, $n = 1, 2, \ldots, N$.

2. Compute the firing interval of the $n^{th}$ rule, $F^n$(**x'**):

$$F^n(\mathbf{x'}) = [\mu_{\underline{X}_1^n}(x'_1)\times\cdots\times\mu_{\underline{X}_I^n}(x'_I), \mu_{\overline{X}_1^n}(x'_1)\times\cdots\times\mu_{\overline{X}_I^n}(x'_I)] \equiv [\underline{f}^n, \overline{f}^n], \quad n = 1, \ldots, N \tag{3.1}$$

   Instead of the product the operator mininum could be used in 3.1. In this research minimum and maximum operators have been used for *AND* and *OR* operators respectively.

3. Perform type-reduction to combine $F^n$(**x'**) and the corresponding rule consequents. The most commonly used method is the center-of-sets type-reducer. [12]

$$Y_{cos}(\mathbf{x'}) = \bigcup_{\substack{f^n\in F^n(\mathbf{x'}) \\ y^n\in Y^n}} \frac{\sum_{n=1}^N f^n y^n}{\sum_{n=1}^N f^n} = [y_l, y_r] \tag{3.2}$$

   It has been show that [12]:

$$y_l = \min_{k\in[1,N-1]} \frac{\sum_{n=1}^k \overline{f}^n \underline{y}^n + \sum_{n=k+1}^N \underline{f}^n \underline{y}^n}{\sum_{n=1}^k \overline{f}^n + \sum_{n=k+1}^N \underline{f}^n} \equiv \frac{\sum_{n=1}^L \overline{f}^n \underline{y}^n + \sum_{n=L+1}^N \underline{f}^n \underline{y}^n}{\sum_{n=1}^L \overline{f}^n + \sum_{n=L+1}^N \underline{f}^n} \tag{3.3}$$

$$y_r = \min_{k\in[1,N-1]} \frac{\sum_{n=1}^k \underline{f}^n \overline{y}^n + \sum_{n=k+1}^N \overline{f}^n \overline{y}^n}{\sum_{n=1}^k \underline{f}^n + \sum_{n=k+1}^N \overline{f}^n} \equiv \frac{\sum_{n=1}^R \underline{f}^n \overline{y}^n + \sum_{n=R+1}^N \overline{f}^n \overline{y}^n}{\sum_{n=1}^R \underline{f}^n + \sum_{n=R+1}^N \overline{f}^n} \tag{3.4}$$

   where the *switch points* L and R are dtermined by

$$\underline{y}^L \le y_l \le \underline{y}^{L+1} \tag{3.5}$$

$$\overline{y}^R \leq y_r \leq \overline{y}^{R+1} \tag{3.6}$$

and $\{\underline{y}^n\}$ and $\{\overline{y}^n\}$ have been sorted in ascending order, respectively. In order to compute $y_l$ and $yr$ Karnik-Mendel (KM) algorithms [12] are performed as follows:

**KM Algorithm for Computing $y_l$**:

(a) Sort $\underline{y}^n$ $(n = 1, 2, \ldots, N)$ in increasing order and call the sorted $\underline{y}^n$ by the same name, but now $\underline{y}^1 \leq \underline{y}^2 \cdots \leq \underline{y}^N$. Match the weights $F^n$(**x'**) with their respective $\underline{y}^n$ and renumber them so that their index corresponds to the renumbered $\underline{y}^n$.

(b) Initialize $f^n$ by setting

$$f^n = \frac{\underline{f}^n + \overline{f}^n}{2} \tag{3.7}$$

and then compute

$$y = \frac{\sum_{n=1}^{N} \underline{y}^n f^n}{\sum_{n=1}^{N} fn} \tag{3.8}$$

(c) Find switch point $k$ $(1 \leq k \leq N-1)$ such that

$$\underline{y}^k \leq y \leq \underline{y}^{k+1} \tag{3.9}$$

(d) Set

$$f^n = \begin{cases} \overline{f}^n & \text{if } n \leq k \\ \underline{f}^n & \text{if } n > k \end{cases} \tag{3.10}$$

and compute

$$y' = \frac{\sum_{n=1}^{N} \underline{y}^n f^n}{\sum_{n=1}^{N} f^n} \tag{3.11}$$

(e) Check if $y' = y$. If yes, stop and set $y_l = y$ and $L = k$. If no, go to Step f).

(f) Set $y = y'$ and go to Step c).

**KM Algorithm for Computing $y_r$**:

(a) Sort $\overline{y}^n$ $(n = 1, 2, \ldots, N)$ in increasing order and call the sorted $\underline{y}^n$ by the same name, but now $\overline{y}^1 \leq \overline{y}^2 \cdots \leq \overline{y}^N$. Match the weights $F^n(\mathbf{x'})$ with their respective $\overline{y}^n$ and renumber them so that their index corresponds to the renumbered $\overline{y}^n$.

(b) Initialize $f^n$ by setting

$$f^n = \frac{\underline{f}^n + \overline{f}^n}{2} \tag{3.12}$$

and then compute

$$y = \frac{\sum_{n=1}^{N} \overline{y}^n f^n}{\sum_{n=1}^{N} fn} \tag{3.13}$$

(c) Find switch point $k$ $(1 \leq k \leq N-1)$ such that

$$\overline{y}^k \leq y \leq \overline{y}^{k+1} \tag{3.14}$$

(d) Set

$$f^n = \begin{cases} \underline{f}^n & \text{if } n \leq k \\ \overline{f}^n & \text{if } n > k \end{cases} \tag{3.15}$$

and compute

$$y' = \frac{\sum_{n=1}^{N} \overline{y}^n f^n}{\sum_{n=1}^{N} f^n} \tag{3.16}$$

(e) Check if $y' = y$. If yes, stop and set $y_l = y$ and $R = k$. If no, go to Step f).

(f) Set $y = y'$ and go to Step c).

4. Compute the defuzzified output as:

$$y = \frac{y_l + y_r}{2} \tag{3.17}$$

[4, 6, 12]

## 3.2 Rough sets

Rough set theory is a new mathematical approach to imperfect knowledge. Rough set theory has found many interesting applications. The rough set approach seems to be of fundamental importance to AI and cognitive sciences, especially in the areas of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems, inductive reasoning and pattern recognition. [13]

In this research rough set theory is used in the same way as in [14], to identify the reduct and the optimal set of decision rules, derived from a decision table.

A rough set is interpreted as a formal approximation of a crisp set, by a pair of sets, which give the so called *lower* and *upper* approximation of the original set. Let consider the classical Pawlak's information system: $IS =_{df} (U, A, V, f)$, where: $U$ is some universe, $A$ is a set of attributes, $V$ is the attributes domain set $V =_{df} \bigcup_{a \in A} V_a$, $V_a$ - is the domain of the $a^{th}$ attribute ($a \in A$) and $f$ is the information function - $f : U \times A \to V$, $\forall_{x \in U, a \in A} f(u, a) \in V_a$. Regarding to the following equivalence relation: $IND(B) =_{df} (x, y) \in U \times U : \forall_{a \in B} f(x, a) = f(y, a)$, where $B \subseteq A$, the lower and the upper approximation of a subset of $U$ can be introduced as follows:

$$B \downarrow X =_{df} \{x \in U : [x]_{IND} \subseteq X\}, \tag{3.18}$$

$$B \uparrow X =_{df} \{x \in U : [x]_{IND} \cap X \neq \varnothing\}, \text{where } X \subseteq U \text{ and } B \subseteq A. \tag{3.19}$$

The above information system can be interpreted and realized as a classical decision table (assuming a decision attribute). Using the mathematical apparatus defined for rough sets, there is possible to identify the reduct and the optimal set of decision rules, derived from a given decision table. The reduct gives the minimal set of attributes that fully characterize the knowledge represented in the equivalence class structure. Next, over the derived reduct, the minimal set of decision rules that cover the corresponding decision problem, can be generated. Also, if there is data with conflict objects - i.e. two objects are conflicting when they are characterized by the same values of all attributes, but they belong to different decision classes, the *lower* and the *upper approximation*

*precision* can be used to eliminate decision table inconsistency. In our work, we have used the rough sets to extract optimal set of decision rules, in the following sequence:

1. Design a decision table, regarding to the considered problem (identify the set of objects, attributes and decision classes).

2. Eliminate object conflicts, using the lower approximation precision: let $X \subseteq U(X \neq \varnothing)$ and $B \subseteq A(B \neq \varnothing)$, then the lower approximation precision of the set X regarding to the ubset B is defined as follows:

$$\gamma_B(X) =_{df} \frac{|B \downarrow X|}{|U|} \tag{3.20}$$

3. Generate the attributes reduct: generate *discernibility matrix* ( M(IS) - discernibility matrix for information system IS: $M(IS) =_{df} [m_{i,j}]_{i,j=1,...n} = \{a \in A : f(x_i, a) \neq f(x_j, a)\}$, where $n = |U|$) and next apply the Johnson heuristic algorithm for rough set reduction, applied to find single reduct (subset of attibutes).

4. Generate the minimal set of decision rules that correctly cover the decision problem (see algorithm below).

*Algorithm 1* (rule extraction):
To generate the optimal set of rules over adecision table, the folowing steps should be taken:
*Step 1*: Generate the 'M$_k$' matrixes (matrixes derived over M(IS), which are used to define the so called implicants of the considered objects) from the discernibility matrix.
*Step 2*: Define the object implicants.
*Step 3*: Define the target set of rules.

Step 1:
($k = 1, ..., n = |U|$) Let $c_{ij}$ are the elements of M(IS), $\hat{c}_{ij}$ are the elements of M$_k$ (with respect to k) and a* is the decision attribute. Then:
*For each k = 1, ..., n*:

1) *If* $i \neq k$ *then* $\hat{c}_{ij} =_{df} \varnothing$

2) *If* $(c_{kj} \neq \varnothing)$ *and* $(d_a^B) * (x_j) \neq \{a^*(x_i)\})$ *then* $\hat{c}_{ij} =_{df} c_{kj} \cap B$ *else* $\hat{c}_{ij} =_{df} \varnothing$,

where $B \subseteq A$ and $d_a^A * (x_k) =_{df} \{v \in V_{a*} | \exists_{y \in U}(x_i \text{ IND } y) \wedge a^*(y) = v\}$

*Example 1:* Let consider the following decision table (extended with the corresponding $d_a^A*$ values; A = B = {a,b,c}):

| Object | a | b | c | a* | $d_a^A*$ |
|---|---|---|---|---|---|
| $x_1$ | 1 | 0 | 1 | 0 | {0} |
| $x_2$ | 0 | 0 | 0 | 1 | {1} |
| $x_3$ | 2 | 0 | 1 | 0 | {0} |
| $x_4$ | 0 | 0 | 1 | 2 | {2} |
| $x_5$ | 1 | 1 | 1 | 0 | {0} |

The corresponding discernibility matrix (*it has the symmetric property)* takes the form:

$$\mathrm{M(SI)} = \begin{pmatrix} \varnothing & \cdots & \cdots & \cdots & \cdots \\ \{a,c\} & \varnothing & \cdots & \cdots & \cdots \\ \{a\} & \{a,c\} & \varnothing & \cdots & \cdots \\ \{a\} & \{c\} & \{a\} & \varnothing & \cdots \\ \{b\} & \{a,b,c\} & \{a,b\} & \{a,b\} & \varnothing \end{pmatrix}$$

Next, applying the *Algorithm 1* for k = 1, we can generate the matrix $M_1$:

$$\mathrm{M}_1 = \begin{pmatrix} \varnothing & \{a,c\} & \varnothing & \{a\} & \varnothing \end{pmatrix}$$

as:

$d_a^A*(x_2), d_a^A*(x_4) \neq \{0\}$ (j=2,4), thus $\hat{c}_{ij} = c_{ij} \cap A(j = 2, 4)$, so : $\{a, c\} \cap A = \{a, c\}, \{a\} \cap A = \{a\}$.

Similarly, we can generate the matrixes: $M_2$, $M_3$, $M_4$, $M_5$.

Step 2:

Next, we can determine the set of *'object implicants'* from each matrix:

*Implicant 1*: from $M_1$ : $x_1 \Rightarrow (a \vee c) \wedge a$ (we can simplify, assuming Boolean algebra and using the corresponding Boolean algebra reduction rules: $x \wedge x = x = x \vee x; x \wedge (x \vee y) = x$ and $x \vee (x \wedge y) = x$) : $x_1 \Rightarrow a$,

*Implicant 2*: from $M_2$: $x_2 \Rightarrow c$,

*Implicant 3*: from $M_3$: $x_3 \Rightarrow a$,

*Implicant 4*: from $M_4$: $x_4 \Rightarrow a \wedge c$,

*Implicant 5*: from $M_5$: $x_5 \Rightarrow a \vee b$.

*(intuitively, the object implicants can be considered as indication concerning which attributes are strongly related to which objects).*

Step 3:

Finally, using the above implicants, we can generate the target set of rules, derived from the considered decision table. Each rule concern one decision value and it is derived as a sum of the object implicants related to that decision, i.e.:

$Rule_1$ : $f(x_1, a^*) = f(x_3, a^*) = f(x_5, a^*) = 0$: $f(x_1, a) \vee f(x_3, a) \vee (f(x_5, a) \vee f(x_5, b)) \Rightarrow (decision : 0)$,

$Rule_2$: $f(x_2, c) \Rightarrow (decision : 1)$,

$Rule_3$: $f(x_4, a) \wedge f(x_4, c) \Rightarrow (decision : 2)$.

# Chapter 4

# Description of the proposed concept

The clue of the decision support method presented in this thesis, is our proposition to combine rough sets and Type-2 fuzzy control, by interpreting the corresponding features, derived from HER-2/neu histopathology images, as domains over which fuzzy sets are defined. This allows to use the fuzzy control concept to generate numerical outputs and then decision values.

However, our system can work in several different kinds of data due that the feature extractor interface is independent from the fuzzyfier and inference interface. This research is mainly focused in demonstrate that the combination of rough sets to obtain the optimal set of rules and Type-2 fuzzy control as a inference mechanism can be used as a decision making support system. This chapter, is divided in two different sections. The first one is independent from the feature extractor interface. We used data from [15] that has been already tested and demonstrated that can be used in classification tasks. The second one, is focused in the application of the whole procedure using HER-2/neu histopathology images as input.

## 4.1 Method applied in UCI dataset

In order the test the effectiveness of rough sets and Type-2 fuzzy control a data set from UCI repository has been chosen. Breast Cancer Wisconsin (Diagnostic) Dataset [16] has

| # | Attribute | Domain |
|---|---|---|
| 1 | Sample Code Number | id number |
| 2 | Clump Thickness | 1-10 |
| 3 | Uniformity of Cell Size | 1-10 |
| 4 | Uniformity of Cell Shape | 1-10 |
| 5 | Marginal Adhesion | 1-10 |
| 6 | Single Epithelial Cell Size | 1-10 |
| 7 | Bare Nuclei | 1-10 |
| 8 | Bland Chromatin | 1-10 |
| 9 | Normal Nucleoli | 1-10 |
| 10 | Mitoses | 1-10 |
| 11 | Class: | 0-bening, 1-malignant |

TABLE 4.1: Wisconsin Dataset Description

been chosen as our dataset due to the similarities (clinical purpose) with our real input data - HER-2/neu histopathology images. The dataset consists in 699 instances with 10 attributes each:

The distribution between the classes is: Benign: 458 instances (65.5%), Malignant: 241 (34.5%).

The dataset has been divided in two: Learning Dataset and Validation Dataset. As explained in Chapter 2 different procedures are applied for each dataset. The main objective with our Learning Dataset is to extract knowledge (in our case fuzzy rules). This fuzzy rules will be used by our inference mechanism to predict the decision value in our validation set.

As shown in 2.1 our input is a crisp dataset. We need to transform this crisp dataset into a fuzzy dataset.

We use the following fuzzification concept: for every feature over the considered learning set, a Gaussian distribution was generated, which was used to define the fuzzy set medium over each feature: $\mu_{medium} = e^{\frac{-(x-x_0)^2}{2\sigma^2}}$, where $x_0$ is the expected value and $\sigma$ is the standard deviation. Next, the fuzzy sets 'low' and 'high' were defined as follows:

$$\mu_{low} = \begin{cases} 1 - e^{\frac{-(x-x_0)^2}{2\sigma^2}} & : \mathrm{x} < \mathrm{x}_0 \\ 0 & : \mathrm{x} \geq \mathrm{x}_0 \end{cases}, \mu_{high} = \begin{cases} 0 & : \mathrm{x} \leq \mathrm{x}_0 \\ 1 - e^{\frac{-(x-x_0)^2}{2\sigma^2}} & : \mathrm{x} > \mathrm{x}_0 \end{cases}$$

For instance, let 1000025 be the Sample Code Number of one of our objects and Thickness(1000025)=5 the value of the attribute Thickness for our object. Then, f(1000025, Thickness)='medium' if and only if $\mu_{medium}(x_{1000025}) > max(\mu_{low}(x_{1000025}), \mu_{high}(x_{1000025}))$.

One of the most important factors in classifying a tumor as benign or malignant is its invasive potential. Therefore, we can consider the decision attributes as fuzzy sets regarding the invasive potential of the tumour. The decision attribute benign will be represented as a fuzzy set that means 'low invasive potential' and the decision attribute malignant will be represented as a fuzzy set meaning 'high invasive potential'. This fuzzy sets will be used in our conclusions using the well-known Logistic function. Therefore, we define the membership of the fuzzyset 'high invasive potential' as $\mu_{highINV} = \frac{1}{1+e^{-x}}$ and the membership of the fuzzy set 'low invasive potential' as $\mu_{lowINV} = 1 - \frac{1}{1+e^{-x}}$. The process for the object confliction elimination, the attributes reduction and the generation of optimal set of rules has been already explained in chapter 3. Next, the generated fuzzy rules were exploited by our inference mechanism. In order to test the improvement that Type-2 fuzzy control can achieve both Mamdani reasoning and Type-2 fuzzy control using KM algorithms explained in previous chapters were implemented and tested.

For better explanation, let consider next example:

*Example:*
Let suppose that we have the following learning dataset:

| Object | $feature_1$ | $feature_2$ | Decision |
|---|---|---|---|
| C1 | medium | large | $D_2$ |
| C2 | small | medium | $D_1$ |
| C3 | large | medium | $D_2$ |

We have three different fuzzy sets: {small, medium, large} defined over the domain of the considered features.

Then we propose the following interpretation of information functions:
$f(C1, feature_1) =' small' \iff$ 'feature of C1 is small.
i.e. $\mu_{small}(x_{C1}) > max\{\mu_{medium}(x_{C1}), \mu_{large}(x_{C1})\}, x \in feature1$.

Next, let suppose that we have generated the following set of decision rules over the decision table:

- $f(C1, feature_1) \vee f(C2, feature_2) \rightarrow$ decision: $D_1$

- $f(C3, feature_1) \wedge f(C3, feature_2) \rightarrow$ decision: $D_2$

Thus, we propose the following interpretation of the above rules in terms of fuzzy rules:

- IF ($feature_1$ is '$medium$') $\oplus$ ($feature_2$ is '$medium$') THEN (decision is $D_1$)

- IF ($feature_1$ is '$large$') $\otimes$ ($feature_2$ is '$medium$') THEN (decision is $D_2$)

where $\otimes$, $\oplus$ are assumed as the Zadeh's triangular norms.

Once, if we derive the set of fuzzy rules that correspond to the considered problem, we are able to apply Type 2 fuzzy logic reasoning.

For better understanding an input example is shown.

The input for the inference mechanism is the validation dataset. The validation data set contains different objects. Same process to give a decision value is applied for all of them. Let consider an example:

| Object | $feature_1$ | $feature_2$ | Decision |
|---|---|---|---|
| T1 | 233 | 26 | $D_2$ |
| T2 | 111 | 233 | $D_1$ |

We need to compute the membership of each value to the corresponding fuzzy sets. Remember that Gaussian distribution with an offset number is used to model our fuzzy sets. Let assume that we get these values for the object T1:

1. $[\mu_{\underline{small}}(feature_1), \mu_{\overline{small}}(feature_1)] = [0.2, 0.3]$
2. $[\mu_{\underline{medium}}(feature_1), \mu_{\overline{medium}}(feature_1)] = [0.45, 0.55]$
3. $[\mu_{\underline{large}}(feature_1), \mu_{\overline{large}}(feature_1)] = [0.9, 1.0]$
4. $[\mu_{\underline{small}}(feature_2), \mu_{\overline{small}}(feature_2)] = [0.77, 0.87]$
5. $[\mu_{\underline{medium}}(feature_2), \mu_{\overline{medium}}(feature_2)] = [0.23, 0.33]$
6. $[\mu_{\underline{large}}(feature_2), \mu_{\overline{large}}(feature_2)] = [0.05, 0.15]$

Then we need to compute the firing intervals of the two rules:

1. $\mathbf{R}_1 \rightarrow [\underline{f}^1, \overline{f}^1] = [max(0.45, 0.23), max(0.55, 0.33)] = [0.45, 0.33]$

2. $\mathbf{R}_2 \rightarrow [\underline{f}^2, \overline{f}^2] = [min(0.9, 0.23), min(1.0, 0.33)] = [0.23,0.33]$

Finally, getting the corresponding interval consequents using the appropiate function -Logit function in our case- we can apply KM algorithms to generate a crisp output.

The experiment in this case was quite simple but enough to show the success of applying Type-2 fuzzy control. The dataset for validation consisted on 233 objects, the rest were used as a learning dataset. With the optimal threshold value the results were 78% in the case of Mamdani Reasoning and 98% in the case of Type-2 control. Note that the theory for Mamdani Reasoning has been avoided in this thesis. For more information about Mamdani check [17] or [14] where the research was focused in this kind of reasoning.

## 4.2 Method applied in HER-2 histopathology images

From the very beginning we wanted to apply the specificated methodology in real clinical data of HER-2/neu breast cancer histopathology images. Therefore, in this case our input is not already a prepared dataset but images. A feature extractor is needed in order to extract to corresponding features of these images. An example of an HER-2/neu breast cancer histopathology is shown in 4.1

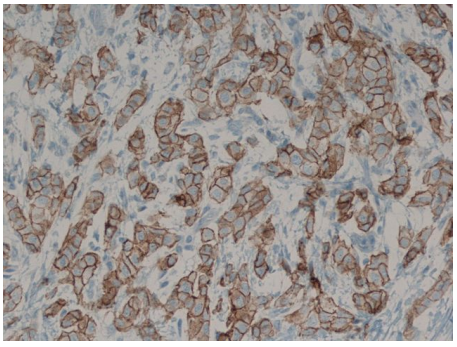


FIGURE 4.1: HER-2/neu histopathology image

Specific cells have been chosen by experts to be analysed (those that were more difficult to determine whether they were aggressive or not). In order to extract the corresponding features of these cells a text file was provided with the coordinates of the cells and all the information needed to determine whether a pixel belonged to a determined cell or not. With all this information the feature extractor is used to build a dataset considering

each cell an object and the features of the cells as attributes. The dataset is composed in the following way:

| Attr name | Definition | Domain |
|---|---|---|
| I | Intensity | [0,255] |
| R_R | Mean Red Value (RGB Model) | [0,255] |
| R_G | Mean Green Value | [0,255] |
| R_B | Mean Blue Value | [0,255] |
| E_R | Mean Amount of Red | $\mathbb{R}$ |
| E_G | Mean Amount of Green | $\mathbb{R}$ |
| E_B | Mean Amount of Blue | $\mathbb{R}$ |
| V | Mean Value Value (HSV Model) | [0,255] |
| S | Mean Saturation Value | [0,255] |
| H | Mean Hue Value | [0,255] |
| *Class | Level of aggressiveness of the cell | {1,2} |

TABLE 4.2: HER-2/neu Dataset Description

The built database consists in more than 1,500 objects (cells), each of them with a decision value that give us the aggressiveness of the cell regarding to HER-2 overexpression. These decision are given by experts, previously diagnosed by FISH examination.

The fuzzification of the considered image features is applied in the same way as in the UCI Dataset. Actually, all the procedure from now and on is the same as in the UCI Dataset. Therefore, we have defined three fuzzy sets: small, medium and large over any feature image. We can consider the decision attributes as fuzzy sets regarding the invasive potential of the tumour.

The conclusions of the rules will be intervals defined by the logistic function (as previously). Let us explain this deeper. In order to have intervals from the domain of our logistic functions following procedure is performed:

- **Computation of the firing intervals for each rule**. The intervals will belong to $[0,1]^2$ due the codomain of our antecedents (Gaussian density function).
- **Logit function.** We compute the logit function with our firing intervals. This will give us the domain of the Logistic function.

$$Y^n = [\underline{y}^n, \overline{y}^n] = [Logit(\underline{f}^n), Logit(\overline{f}^n)]$$

Fig 4.2 show logit and logistic function respectively. Blue plot represents the conclusion 'high aggressiveness' and red plot represents 'low aggressiveness' conclusion.

Once we have our intervals defined KM algorithms explained in Chapter 3 are used to compute an output.

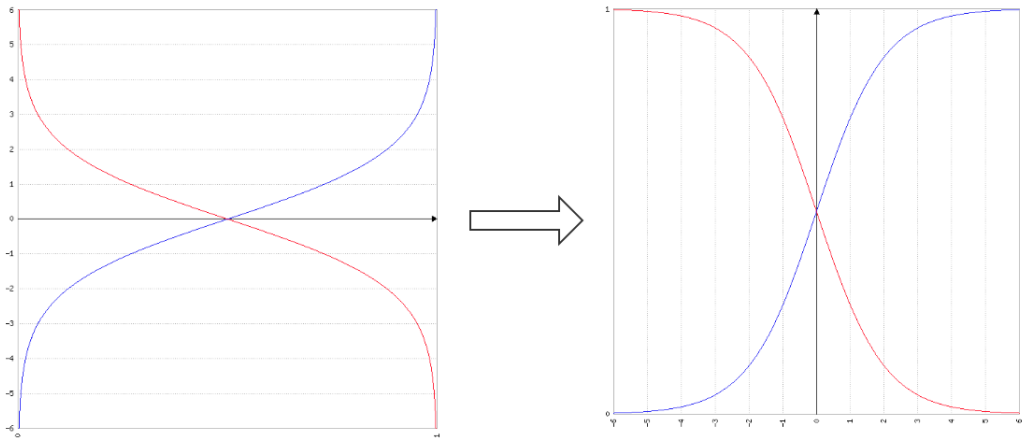


FIGURE 4.2: Logit and Logistic functions

These values can be interpretable as a decision making support. The way we have defined the conclusions mean that high output values refer to high cancer HER overexpression and thus appropriate for trastuzumab treatment. The HER2 overexpression is related (in terms of histopathology HER2 image information) to darker cell membrane staining.

So, the HER2 overexpresion is related to high system output values, which is very important information for physicians. This can be also used for decision making in quantitative manner, i.e. *IF* most of the cell values of a considered HER2 image are $\leq$ *T* (where T is a priori given threshold value), *THEN* classify as appropriate for trastuzumab treatment *ELSE* classify as not appropriate for trastuzumab treatment.

# Chapter 5

# Experiments and results

In order to test the accuracy of our system we have tried to predict the aggressiveness of the cells. To do this, a reasonable threshold value has been chosen to differentiate the output between different classes, i.e. *IF* output $\leq T \implies$ aggressiveness is low, *ELSE* $\implies$ aggressiveness is high. Some results for specific marked cells are shown in 5.1. It is clear that a threshold value can be used to differentiate between the two different classes.

A simple cross validation has been used, concretely 5-fold validation. The original sample was partitioned into 5 equal sized subsamples. Of the 5 subsamples, a single subsample is retained as the validation data for testing the model, and the remaining 4 subsamples were used as training data. The cross-validation process is then repeated 5 times (the folds), with each of the 5 subsamples used exactly once as the validation data. The results achieved are shown in 5.2.

It should be noticed that different parameters are possible to change in case of Type-2 fuzzy control. In our case we have used an *offset* (for the membership of Type-2 fuzzy sets) and a *threshold* value chosen by us, i.e. not optimization method has been involved to test optimal *offset* and *threshold* values. Furthermore, different functions in antecedents and in consequents could be implemented. In this case, we have generalized our image features using Gaussian density functions and our conclusions using Logistic functions (Gaussian in Mamdani conclusions, Logistic in type-2) but this is not a constraint, some other functions can be tested.

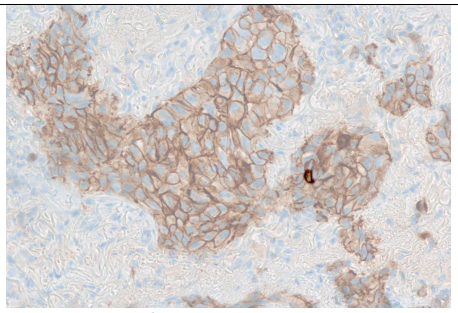
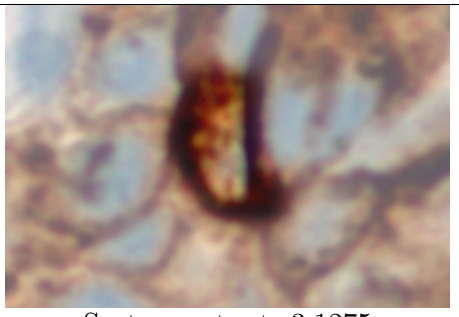
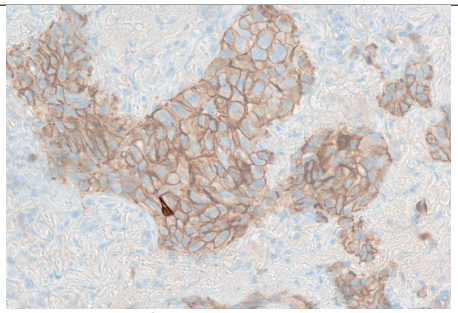
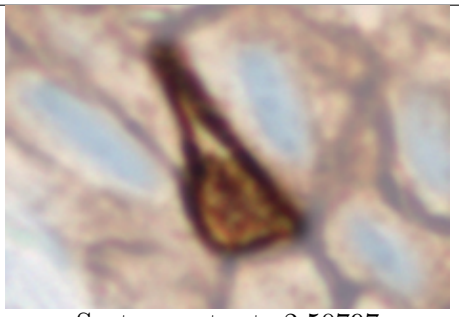
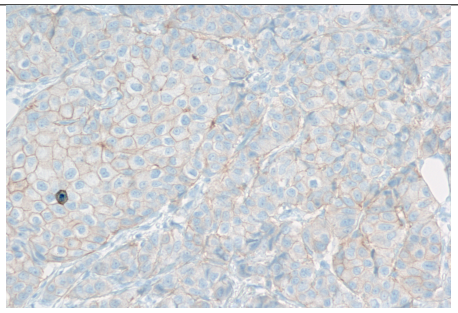
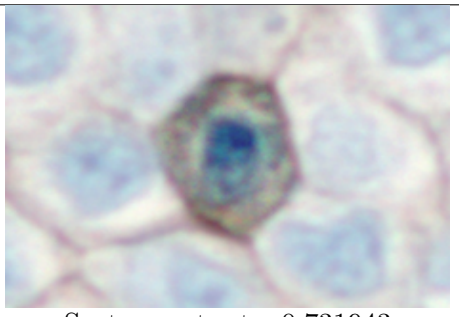
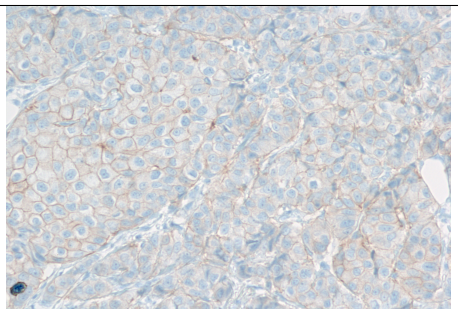
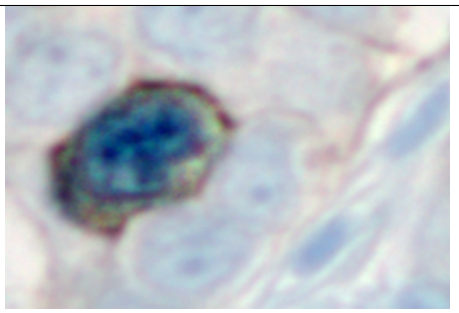
| Histopathology image and FISH diagnose | Marked cell zoom and system output |
|---|---|
|  High aggressiveness |  System output: 3.1275 |
|  High aggressiveness |  System output: 2.50797 |
|  Low aggressiveness |  System output: -0.731943 |
|  Low aggressiveness |  System output: -0.976381 |

TABLE 5.1: Cells classification

## 5.1 Parameters of Logistic Function

Previous experiments were performed with the most basic Logistic Function: $f(x) = \frac{1}{1+e^{-t}}$. However, this function can be extended using two new parameters:

$$f(x) = \frac{1}{1 + e^{-k(x-x_0)}} \tag{5.1}$$

where $x_0$ is the x-value of the sigmoid's midpoint and $k$ is the steepness of the curve.

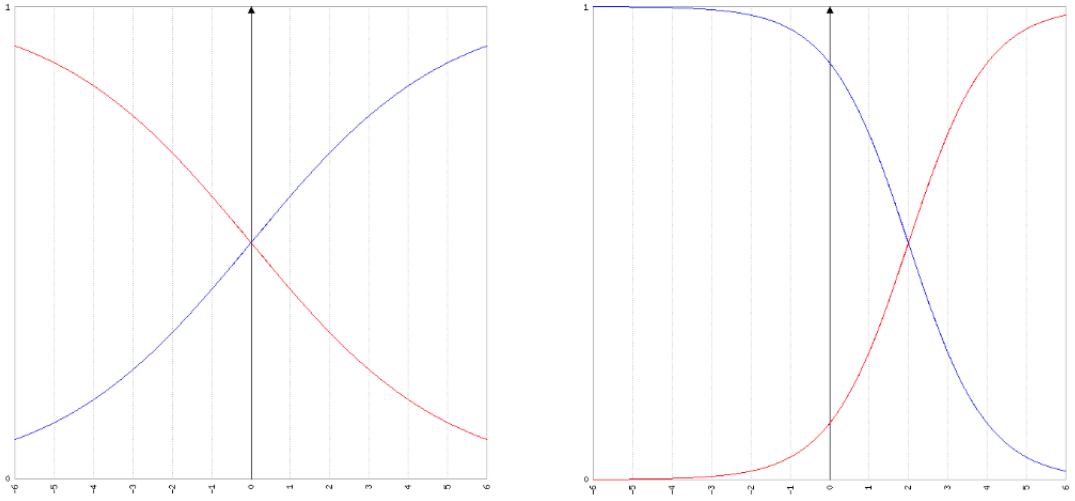


FIGURE 5.1: Steepness and midpoint parameters

Figure 5.1 shows how different parameters for Logistic Function. Notice how steepness change the slope of the curve, it makes it smoother the curve while decreasing the value. And, the parameter midpoint change the crossing point in the x-axis between the functions.

Several experiments with different parameters were performed. In order to automatise the process the threshold value to decide whether a cell is aggressive or not is automatically calculated as: $T = \frac{(\overline{y}_0 + \overline{y}_2)}{2}$, where $\overline{y}_0$ and $\overline{y}_2$ are the mean of the defuzzified values in decision value 0 and 2 respectively. The dataset partition used for this experiment is the #1 in Table 5.2. The results are shown in Figs. 5.2 and 5.3.

It has been shown that the accuracy of the system can be extremely different with different parameters. Therefore, an optimization method to find the optimal crossing

| # | Mamdani Reasoning (%) | Type-2 fuzzy control (%) |
|---|---|---|
| 1 | 75.37 | 74.16 |
| 2 | 77.20 | 74.77 |
| 3 | 76.29 | 74.77 |
| 4 | 65.88 | 65.80 |
| 5 | 65.65 | 67.17 |

TABLE 5.2: System accuracy

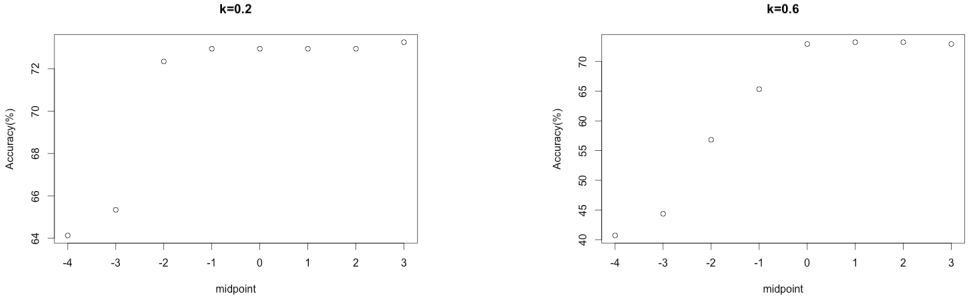



FIGURE 5.2: Results with k=0.2 and k=0.6

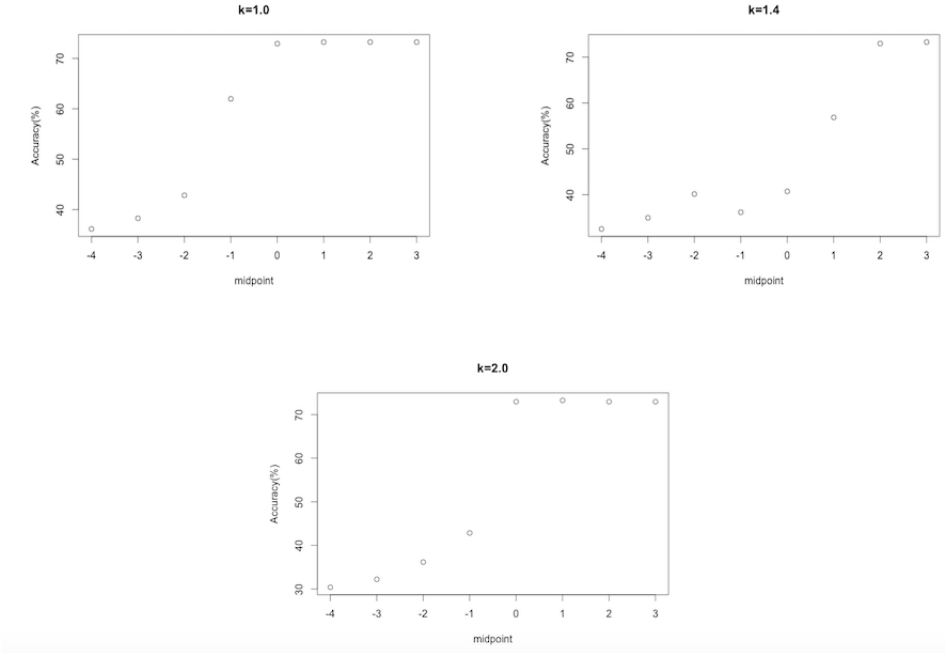



FIGURE 5.3: Results with k=1.0, k=1.4 and k=2.0

point and the optimal steepness of the curve could arise the accuracy of the system.

# Chapter 6

# Conclusions and further research

The contemporary for breast cancer diagnosis is visual examination of microscopic biopsy images. Nevertheless, manual expert classification is both inefficient and ineffective, hence more robust and automatic approach is needed. In this thesis, an histopathology decision making support system, based on rough sets and type-2 fuzzy control for HER-2/neu image analysis is presented.

In presented solution, preprocess phase enables to focus on the most crucial parts of biopsy image, the cells themselves. Applying fuzzy sets to features vector helps to reflect fuzzy characteristic of image attributes, providing more appropriate description then crisp value. The rough sets successfully helped to reduce features vector and remove conflicting objects from dataset, providing foundation for more efficient classification. Lastly, fuzzy rules and fuzzy control system along with Mamdani reasoning leads to precise classification. System accuracy is estimated as 70% concordance with conclusive FISH test. However, our results with UCI dataset were around 98% which means that the approach is good but the current vector features is not enough good.

There are several ways to improve the presented work. Mainly, a better feature vector should be provided, i.e. improving the feature extractor interface. Currently it only give us colour-related attributes. There are many other features important to decide the HER2 overexpression of a cell such as shape related attributes, texture analysis, etc. Secondly, we have used a narrow range of values for offset and threshold parameters and we selected specific functions. A further research in what functions are more appropiate to use both in rules' antecedents and consequents would be useful as well as the use

of optimization methods to find optimal parameters. Finally, the dataset used in this research should be enlarged to provide more accurate results, which would be essential to test presented approach before using the system in clinical settings.

# Bibliography


[1] Robert A. Greenes Mark A. Musen, Blackford Middleton. Clinical decision-support systems. 72:643–674, 2014. URL http://eknygos.lsmuni.lt/springer/56/698-736.pdf.

[2] Timothy J. Moynihan. Her2-positive breast cancer: What is it? May 2015. URL http://www.mayoclinic.org/breast-cancer/expert-answers/faq-20058066.

[3] Dongrui Wu. A brief tutorial on interval type-2 fuzzy sets and systems. 62, 2007. URL https://sites.google.com/site/drwu09/publications/A_Brief_Tutorial_on_Interval_Type-2_Fuzzy_Sets_and_Systems.pdf?attredirects=0.

[4] Jerry M. Mendel. Type-2 fuzzy sets and systems: How to learn about them. page 8, June 2009. URL http://sipi.usc.edu/~mendel/publications/MendelSMCeNewsletter6-09.pdf.

[5] L. A. Zadeh. Information and control. 8:338–353, June 1965. URL http://www.cs.berkeley.edu/~zadeh/papers/Fuzzy%20Sets-Information%20and%20Control-1965.pdf.

[6] Feilong Liu Jerry M. Mendel, Robert I. John. Interval type-2 fuzzy logic systems made simple. *IEEE TRANSACTIONS ON FUZZY SYSTEMS*, 14: 808–821, December 2006. URL http://sipi.usc.edu/~mendel/publications/MendelJLTFS12-06.pdf.

[7] Zhi Liu and Han-Xiong Li. A probabilistic fuzzy logic system for modeling and control. *Fuzzy Systems, IEEE Transactions*, 13:848 – 859, December 2005. URL http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=1556589&url.

[8] Jerry M. Mendel. Fuzzy logic systems for engineering: A tutorial. *IEEE*, 62:345–377, 1995. URL http://repository.ksu.edu.sa/jspui/bitstream/123456789/2294/1/research_12.pdf.

[9] Jerry M. Mendel. Type-2 fuzzy sets made simple. *Fuzzy Systems, IEEE Transactions*, 10:117–127, April 2002. URL http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=995115&url=http.

[10] Nilesh N. Karnik and Jerry M. Mendel. Operation on type-2 fuzzy sets. *Fuzzy Systems, IEEE Transactions*, 122:327–348, June 2001. URL http://www.sciencedirect.com/science/article/pii/S0165011400000798.

[11] Jerry M. Mendel. Type-2 fuzzy sets and systems: an overview. *Computational Intelligence Magazine, IEEE*, 2:20–29, June 2007. URL http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=4197699&url.

[12] Jerry M. Mendel. Uncertain rule-based fuzzy logic systems: Introduction and new directions. *Prentice Hall*, 2, January 2001.

[13] Z. Pawlak. Rough sets. *International Journal of Computer and Information Sciences*, 11:341–356, October 1982. URL http://link.springer.com/article/10.1007/BF01001956.

[14] Marzenna Podhorska-Okolow Bartosz Pula Jedrzej Grzegrzolka Martin Tabakov, Krzysztof Rodak. Decision support system for histopathological diagnosis of her2 breast cancer using pawlaks information system and mamadani type fuzzy control. *Journal of Intelligent Computing*, 4:30–42, March 2013. URL http://www.dline.info/jic/fulltext/v4n1/5.pdf.

[15] M. Lichman. UCI machine learning repository, 2013. URL http://archive.ics.uci.edu/ml.

[16] Olvi L. Mangasarian Dr. William H. Wolberg, W. Nick Street. Breast cancer wisconsin (diagnostic) data set. URL http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/.

[17] E.H. Mamdani and B. R. Gaines. *Fuzzy Reasonings and Its Applications*.