

Associative Memory Approach for the Diagnosis of Parkinson's Disease

Elena Acevedo^{*}, Antonio Acevedo, and Federico Felipe

Escuela Superior de Ingeniería Mecánica y Eléctrica, IPN, Mexico City, Mexico
{eacevedo,macevedo,ffelipe}@ipn.mx

Abstract. A method for diagnosing Parkinson's disease is presented. The proposal is based on associative approach, and we used this method for classifying patients with Parkinson's disease and those who are completely healthy. In particular, Alpha-Beta Bidirectional Associative Memory is used together with the modified Johnson-Möbius codification in order to deal with mixed noise. We use three methods for testing the performance of our method: Leave-One-Out, Hold-Out and K-fold Cross Validation and the average obtained was of 97.17%.

Keywords: Classification, Associative Models, Alpha-Beta BAM, Codification.

1 Introduction

Parkinson's disease (PD) was first described in a medical context in 1817 by James Parkinson, a general practitioner in London [1]. PD is the second most common neurodegenerative disorder after Alzheimer's disease. It has been suggested that the prevalence of the disease will double over the next 20 years [2].

Parkinson's disease is not an infection but a disease of the brain [3]. It is a chronic condition, an imbalance resulting from a loss of dopamine.

There are four cardinal features of PD [4] that can be grouped under the acronym TRAP: Tremor at rest, Rigidity, Akinesia (or bradykinesia) and Postural instability. The signs of Parkinsonism [5] are showed in Table 1.

There are treatments which help the patient to control PD in order to have a better way of life. But when there is not a previous diagnostic, the patient can suffer the consequences of the disease at an advanced stage, and in most of the time this can mean death. It is assessed that there is a 30% of patients without a diagnostic. However, PD is diagnosed on clinical criteria; there is no definitive test for diagnosis [4]. Certain tests may be done to help diagnose other conditions with similar symptoms. For instance, blood tests may be done to check for abnormal thyroid hormone levels or liver damage. An imaging test (such as a CT scan or an MRI) may be used to check for signs of a stroke or brain tumor.

^{*} María Elena Acevedo Mosqueda, Escuela Superior de Ingeniería Mecánica y Eléctrica del Instituto Politécnico Nacional, Av. IPN s/n Edif Z-4, 3er piso, Telecomunicaciones, Col. Lindavista, C.P. 07738, Mexico City, Mexico. E-mail: eacevedo@ipn.mx.

Table 1. Signs of Parkinsonism

Location or Activity	Manifestation
Face	<ul style="list-style-type: none"> • Loss of animation (masking) • Decreased blink rate
Speech	<ul style="list-style-type: none"> • Reduced volume • Dysarthric due to reduced amplitude and precision of the articulators of speech (lips, tongue, and palate)
Automatic movements	<ul style="list-style-type: none"> • Less gesturing when talking • Reduced arm swing while walking
Gait	<ul style="list-style-type: none"> • May have difficulty rising from seated position • Stooped • Shortened stride • Shuffling of feet (or feet more parallel to floor in contrast to normal landing on heel and pushing off with toes) • May exhibit freezing in place or hesitancy • Takes several steps to turn • Slowness (bradykinesia)
Balance	<ul style="list-style-type: none"> • Imbalance typically not an early sign in Parkinson's disease but may be in other parkinsonian disorders • Pull test may detect milder degrees of imbalance (retropulsion)
Tremor	<ul style="list-style-type: none"> • Hands when in position of repose (e.g., in lap or at sides during walking) • Legs when patient is seated with feet on floor • Chin • Markedly reduced with action (except with concurrent essential tremor)
Tone	<ul style="list-style-type: none"> • Rigidity of limbs and sometimes neck • If superimposed tremor, examiner appreciates cogwheel pattern
Rapid alternating movements	<p>Slowed, reduced amplitude, and sometimes freezing of movement:</p> <ul style="list-style-type: none"> ▪ Finger-thumb tapping ▪ Pronation-supination ▪ Opening-closing fist ▪ Foot or heel tapping
Other	<ul style="list-style-type: none"> • Eye movements may be slowed, and eye movement falls short of target (hypometric) • Meyerson's sign

Another type of imaging test, called PET (Position Emission Tomography), sometimes may detect low levels of dopamine in the brain. However, PET scanning is not commonly used to evaluate Parkinson's disease because it is very expensive, is not available in many hospitals, and is only used experimentally.

Speech analysis is an alternative for diagnosing PD. Speech is the most complex of innately acquired human motor skills [5], an activity characterized in normal adults by the production of about 14 distinguishable sounds per second through the coordinated actions of about 100 muscles innervated by multiple cranial and spinal nerves. The

ease with which we speak belies the complexity of the act, and that complexity may help explain why speech can be exquisitely sensitive to nervous system disease. In fact, changes in speech can be the only evidence of neurologic disease early in its evolution and sometimes the only significant impairment in a progressive or chronic neurologic condition. In such contexts, recognizing the meaning of specific speech signs and symptoms can provide important clues about the underlying pathophysiology and localization of neurologic disease.

In the other hand, a number of rating scales are used for the evaluation of motor impairment and disability in patients with PD [4], but most of these scales have not been fully evaluated for validity and reliability. The Hoehn and Yahr scale is commonly used to compare groups of patients and to provide gross assessment of disease progression, ranging from stage 0 (no signs of disease) to stage 5 (wheelchair bound or bedridden unless assisted). The Unified Parkinson's Disease Rating scale (UPDRS) [6] is the most well established scale for assessing disability and impairment. Studies making use of UPDRS to track the progression of PD suggest that the course of PD is not linear and that the rate of deterioration is variable and more rapid in the early phase of the disease and in patients with the postural instability gait difficulty (PIGD) of PD.

According to the World Health Organization [7] there are six millions of people affected by this disease in the world and fifty thousand in Mexico. The National Parkinson Foundation (NPF) [8] said that there are between 50 and 60 thousand new cases every year. Around de world, the incidence is from 20 to 25 new cases per year for every 100,000 citizens. A 2% from the affected people suffers the disease due to hereditary factors. Therefore, it is important to have the necessary means to classify Parkinson's patients.

Artificial Intelligence (AI) is an area which is extensively used for the classifying task. In medical diagnosis, a suitable classifier could be useful for an expert to increase the accuracy and reliability of the diagnostic and to minimize the possible errors.

Several works have been developed for diagnosing Parkinson's disease. Some of them use PET (Position Emission Tomography) or SPECT (Single Photon Emission Tomography) images as training data for Neural Network based systems [9] and Support Vector Machine based systems [10]. Another work analyzed the algorithms which model kinetics of dopamine by using the Laplace transformation of differential equations and by algebraic computation with the aid of Gröbner base constructions [11]. With the use of this method they obtained a rigorous solution with respect to the kinetic constants over the Laplace domain. Keijers et al. [12] use neural networks for the classification and rating dyskinesia as well as for extracting the important parameters to distinguish between dyskinesia and normal voluntary movements. An algorithm which combines a perceptron neural network with simple signal processing and rule-based classification [13] is used for automatic recognition and classification of walking patterns, in order to recognize disturbances during walking in PD patients.

Other related works used as training data the dataset introduced by Tsanas and Little [14] from the Oxford University. One approach for the classification of patients with Parkinson's disease is Neural Networks [15,16], another used Support Vector Machine algorithm [17], while Gil [18] applied a combination of both approaches.

In this work, we use this dataset but Associative Models are applied as an alternative approach for classifying patients with Parkinson's disease. In particular,

we used the Alpha-Beta Bidirectional Associative Memory [19]. The main feature of this model is its correct recall. It does not present the forgetting factor; every pattern trained is correctly recalled. The algorithm of the model is not an iterative process and has not stability problems. The correct recall is showed no matter the nature of the patterns.

In section 2, we present the basic concepts of Alpha-Beta associative memories, and we introduce the modified Johnson-Möbius code for avoiding mixed noise. Section 3 describes the main model used in this work. We present the results in section 4. Finally, conclusions are presented.

2 Alpha-Beta Associative Memories

An Associative Memory (AM) M is a system that relates input patterns, and outputs patterns. Two phases comprise the design of an AM: learning phase and recalling phase. In the learning phase, the memory is trained by associating input patterns x and output patterns y (see figure 1). Both input and output patterns can represent any association, for example: fingerprints with faces, names with telephone numbers, DNA sequences with names, etc.

Learning Phase

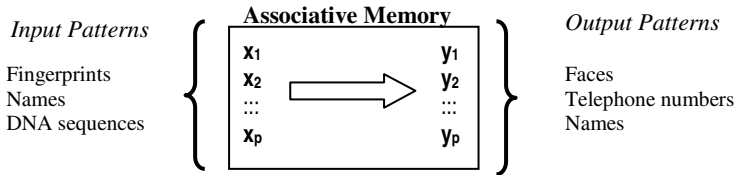


Fig. 1. The learning phase for an Associative Memory

After the associative memory was trained, output patterns can be recalled by presenting the input patterns to the memory. This task is performed by the recalling phase (see figure 2).

Recalling phase



Fig. 2. The recalling phase for an Associative Memory

In figure 2, one can observe that when an input pattern x^k is presented to the AM its corresponding pattern y^k must be recalled. Moreover, if a noisy version of an input pattern x^k represented by \tilde{x}^k is presented to the associative memory, the corresponding pattern y^k should be recalled, if this happens then AM has a correct recall.

Formally, we can say that for a k integer and positive, the corresponding association will be denoted as (x^k, y^k) . The associative memory \mathbf{M} is represented by a matrix whose ij -th component is m_{ij} . Memory \mathbf{M} is generated from an a priori finite set of known associations, known as the fundamental or training set of associations.

If μ is an index, the fundamental set is represented as: $\{(x^\mu, y^\mu) | \mu = 1, 2, \dots, p\}$ with p the cardinality of the set. The patterns that form the fundamental set are called fundamental patterns. If it holds that $x^\mu = y^\mu, \forall \mu \in \{1, 2, \dots, p\}$, \mathbf{M} is *autoassociative*, otherwise it is *heteroassociative*; in this case it is possible to establish that $\exists \mu \in \{1, 2, \dots, p\}$ for which $x^\mu \neq y^\mu$.

2.1 Alpha-Beta Associative Memories

Among the variety of associative memory models described in the scientific literature, there are two models that, because of their relevance, it is important to emphasize: morphological associative memories which were introduced by Ritter *et al* [20], and Alpha-Beta associative memories. Because of their excellent characteristics, which allow them to be superior in many aspects to other models for associative memories, morphological associative memories served as starter point for the creation and development of the Alpha-Beta associative memory.

The Alpha-Beta associative memories [21] are of two kinds (*max* and *min*-type memories) and are able to operate in two different modes. The operator α is useful at the learning phase, and the operator β is the basis for the pattern recall phase. The heart of the mathematical tools used in the Alpha-Beta model, are two binary operators designed specifically for these memories. These operators are defined as follows: first, we define the sets $A=\{0,1\}$ and $B=\{0,1,2\}$, then the operators α and β are defined in Tables 2 and 3, respectively:

Table 2. Alpha operator, $\alpha: A \times A \rightarrow B$

x	y	$\alpha(x,y)$
0	0	1
0	1	0
1	0	2
1	1	1

Table 3. Beta operator, $\beta: B \times A \rightarrow A$

x	y	$\beta(x,y)$
0	0	0
0	1	0
1	0	0
1	1	1
2	0	1
2	1	1

The sets A and B , the α and β operators, along with the usual \wedge (minimum) and \vee (maximum) operators, form the algebraic system $(A, B, \alpha, \beta, \wedge, \vee)$ which is the mathematical basis for the Alpha-Beta associative memories.

We present the learning and recalling phases for an autoassociative memory because the proposed model is designed with this type of memory.

Learning Phase

Step 1. For each $\mu = 1, 2, \dots, p$, from the pair $(\mathbf{x}^\mu, \mathbf{x}^\mu)$ a matrix is built

$$[\mathbf{x}^\mu \otimes (\mathbf{x}^\mu)^t]_{n \times n}$$

Step 2. If the memory is *max*-type, the maximum operator \vee is applied to the matrices obtained in step 1, therefore a *max* matrix \mathbf{V} is built. On the other hand, if the memory is *min*-type then the minimum \wedge operator is applied for building a *min* matrix $\mathbf{\Lambda}$.

$$\mathbf{V} = \vee_{\mu=1}^p [\mathbf{x}^\mu \otimes (\mathbf{x}^\mu)^t] \quad \mathbf{\Lambda} = \wedge_{\mu=1}^p [\mathbf{x}^\mu \otimes (\mathbf{x}^\mu)^t]$$

Recalling Phase

The goal of this phase is recovering the output pattern \mathbf{y}^ω from input pattern \mathbf{x}^ω which is presented to the associative memory. β operator is used in this phase. If at learning phase a *max*-type memory \mathbf{V} was built then pattern \mathbf{x}^ω will operate with matrix \mathbf{V} and the *min* operator \wedge . Pattern \mathbf{x}^ω will operate with *min*-type memory $\mathbf{\Lambda}$ and the *max* operator \vee if a *min*-type memory was built in learning phase.

$$(\vee \Delta_\beta \mathbf{x}^\omega)_i = \wedge_{j=1}^n \beta(v_{ij}, x_j^\omega) \quad \text{or} \quad (\wedge \nabla_\beta \mathbf{x}^\omega)_i = \vee_{j=1}^n \beta(\lambda_{ij}, x_j^\omega)$$

Types of Noise

An associative memory that works with binary values can encounter three types of noise: additive, subtractive and mixed. *Max* and *min*-type Alpha-Beta memories can handle with additive and subtractive noise, respectively. However, none of the two types can handle with mixed noise (see figure 3). Therefore, the patterns need a pre-processing in order to avoid the mixed noise. One way to avoid mixed noise is codifying the patterns using a code which allows the change of one bit. In the following section Johnson-Möbius code is presented.

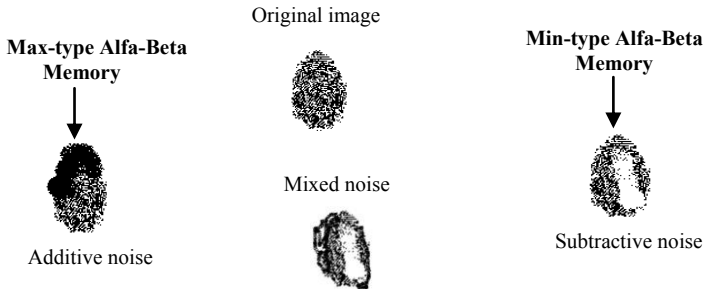


Fig. 3. Types of noise: additive, subtractive and mixed. They could appear when we work with binary values. *Max*-type Alpha-Beta associative memories can handle additive noise while *min*-type memories can handle subtractive noise but none of them handle mixed noise.

2.2 Johnson-Möbius Modified Code

Johnson-Möbius code is a binary code which allows the change of one bit between two consecutive numbers [22]. For representing a decimal number n with Johnson-Möbius code, $n/2$ bits are needed if n is even and $(n+1)/2$ bits are needed if n is odd. The following example shows the algorithm for the Johnson-Möbius code.

Johnson-Möbius Algorithm

Let be the set $r = \{2.5, 0.15, -0.1, 0.4, 1.4\} \subset \mathbf{R}$

Step 1. We can observe that in the set there is a negative number (-0.1), then the set has to be turn into a new set by adding 0.1 to each element. The new set is $t = \{2.6, 0.25, 0.0, 0.5, 1.5\}$

Step 2. Choose a fixed number d of decimals and truncate each of the numbers of the new set to d decimals. The number d depends on the required accurateness for the specific problem we are resolving. In this example we will use $d=1$ then we obtain $e = \{26, 2, 0, 5, 15\}$ where the maximum is $e_m = 26$.

Step 3. As 26 is an even number it is not necessary to add 1.

Step 4. $e_m / 2 = 13$.

Step 5. If $e_i < e_m/2$, then $e_m / 2 - e_i$ zeros are generated and e_i ones are added. In the contrary, $e_m - e_i$ ones are generated and $e_i - e_m/2$ zeros are added. Table 4 shows the results.

Table 4. Johnson-Möbius codification for the set $\{2.5, 0.15, -0.1, 0.4, 1.4\}$

Decimal	Johnson-Möbius codification
26	11111111111111
2	110000000000
0	000000000000
5	111110000000
15	001111111111

Johnson-Möbius Modified Algorithm [23]

N bits are necessary for representing a decimal number n with the Johnson-Möbius code. We present an illustrative example.

Let be the set $r = \{2.5, 0.15, -0.1, 0.4, 1.4\} \subset \mathbf{R}$

Step 1 and Step 2. The same process is performed.

Step 3. For each e_i from the set e , $e_m - e_i$ zeros are concatenated with e_i ones. Table 5 shows the results.

Table 5. Modified Johnson-Möbius codification for the set $\{2.5, 0.15, -0.1, 0.4, 1.4\}$

Decimal	Johnson-Möbius codification
26	111111111111111111111111111111
2	0000000000000000000000000011
0	0000000000000000000000000000
5	00000000000000000000000011111
15	0000000000111111111111111111

Even though more bits are needed for representing a decimal number with Johnson-Möbius modified code, this pre-processing has shown to be better than the original code for avoiding mixed noise.

3 Alpha-Beta Bidirectional Associative Memories

Generally, any bidirectional associative memory model appearing in current scientific literature could be draw as figure 4 shows.



Fig. 4. General scheme of a Bidirectional Associative Memory

General BAM is a “black box” operating in the next way: given a pattern x , associated pattern y is obtained, and given the pattern y , associated pattern x is recalled. Besides, if we assume that \tilde{x} and \tilde{y} are noisy versions of x and y , respectively, it is expected that BAM could recover all corresponding free noise patterns x and y .

Before going into detail over the processing of an Alpha-Beta BAM, we will define the following.

In this work we will assume that Alpha-Beta associative memories have a fundamental set denoted by $\{(\mathbf{x}^\mu, \mathbf{y}^\mu) \mid \mu = 1, 2, \dots, p\}$ $\mathbf{x}^\mu \in A^n$ and $\mathbf{y}^\mu \in A^m$, with $A = \{0, 1\}$, $n \in \mathbb{Z}^+$, $p \in \mathbb{Z}^+$, $m \in \mathbb{Z}^+$ and $1 < p \leq \min(2^n, 2^m)$. Also, it holds that all input patterns are different; M that is $\mathbf{x}^\mu = \mathbf{x}^\xi$ if and only if $\mu = \xi$. If $\forall \mu \in \{1, 2, \dots, p\}$ it holds that $\mathbf{x}^\mu = \mathbf{y}^\mu$, the Alpha-Beta memory will be *autoassociative*; if on the contrary, the former affirmation is negative, that is $\exists \mu \in \{1, 2, \dots, p\}$ for which it holds that $\mathbf{x}^\mu \neq \mathbf{y}^\mu$, then the Alpha-Beta memory will be *heteroassociative*.

Definition 1 (One-Hot). Let the set A be $A = \{0, 1\}$ and $p \in \mathbb{Z}^+$, $p > 1$, $k \in \mathbb{Z}^+$, such that $1 \leq k \leq p$. The k -th one-hot vector of p bits is defined as vector $\mathbf{h}^k \in A^p$ for which it holds that the k -th component is $h_k^k = 1$ and the set of the components are $h_j^k = 0$, $\forall j \neq k, 1 \leq j \leq p$.

Remark 1 In this definition, the value $p = 1$ is excluded since a one-hot vector of dimension 1, given its essence, has no reason to be.

Definition 2 (Zero-Hot). Let the set A be $A = \{0, 1\}$ and $p \in \mathbb{Z}^+$, $p > 1$, $k \in \mathbb{Z}^+$, such that $1 \leq k \leq p$. The k -th zero-hot vector of p bits is defined as vector $\bar{\mathbf{h}}^k \in A^p$ for which it holds that the k -th component is $h_k^k = 0$ and the set of the components are $h_j^k = 1$, $\forall j \neq k, 1 \leq j \leq p$.

Remark 2. In this definition, the value $p = 1$ is excluded since a zero-hot vector of dimension 1, given its essence, has no reason to be.

Definition 3 (Expansion vectorial transform). Let the set A be $A = \{0, 1\}$ and $n \in \mathbf{Z}^+, y, m \in \mathbf{Z}^+$. Given two arbitrary vectors $\mathbf{x} \in A^n$ and $\mathbf{e} \in A^m$, the expansion vectorial transform of order m , $\tau^e : A^n \rightarrow A^{n+m}$, is defined as $\tau^e(\mathbf{x}, \mathbf{e}) = \mathbf{X} \in A^{n+m}$, a vector whose components are: $X_i = x_i$ for $1 \leq i \leq n$ and $X_{i+n} = e_i$ for $n+1 \leq i \leq n+m$.

Definition 4 (Contraction vectorial transform). Let the set A be $A = \{0, 1\}$ and $n \in \mathbf{Z}^+, y, m \in \mathbf{Z}^+$ such that $1 \leq m < n$. Given one arbitrary vector $\mathbf{X} \in A^{n+m}$, the contraction vectorial transform of order m , $\tau^c : A^{n+m} \rightarrow A^n$, is defined as $\tau^c(\mathbf{X}, m) = \mathbf{c} \in A^n$, a vector whose components are: $c_i = X_{i+n}$ for $1 \leq i < m$.

In both directions, the model is made up by two stages, as shown in figure 5. For simplicity, we will just described the process necessary in one direction, $x \rightarrow y$.

The function of Stage 2 is to offer a \mathbf{y}^k as output ($k = 1, \dots, p$) given a \mathbf{x}^k as input. Now we assume that as input to Stage 2 we have one element of a set of p orthonormal vectors. Recall that the *Linear Associator* has perfect recall when it works with orthonormal vectors. In this work we use a variation of the *Linear Associator* in order to obtain \mathbf{y}^k from a one-hot vector \mathbf{h}^k in its k -th coordinate.

For the construction of the modified Linear Associator, its learning phase is skipped and a matrix \mathbf{M} representing the memory is built. Each column in this matrix corresponds to each output pattern \mathbf{y}^μ . In this way, when matrix \mathbf{M} is operated with a one-hot vector \mathbf{h}^k , the corresponding \mathbf{y}^k will always be recalled.

The task of Stage 1 is: given a \mathbf{x}^k or a noisy version of it ($\tilde{\mathbf{x}}^k$), the one-hot vector \mathbf{h}^k must be obtained without ambiguity and with no condition. In its learning phase, stage 1 has the following algorithm:

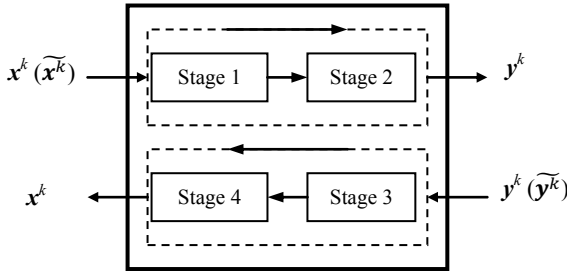


Fig. 5. Alpha-Beta BAM model scheme

1. For $1 \leq k \leq p$ do expansion: $\mathbf{X}^k = \tau^e(\mathbf{x}^k, \mathbf{h}^k)$
2. For $1 \leq i \leq n$ and $1 \leq j \leq n$: $v_{ij} = \sqrt{\alpha} \prod_{\mu=1}^p (X_i^\mu, X_j^\mu)$
3. For $1 \leq k \leq p$ do expansion: $\bar{\mathbf{X}}^k = \tau^c(\mathbf{x}^k, \bar{\mathbf{h}}^k)$

4. For $1 \leq i \leq n$ and $1 \leq j \leq n$:

$$\lambda_{ij} = \bigwedge_{\mu=1}^p \alpha(\bar{X}_i^\mu, \bar{X}_j^\mu)$$

5. Create modified Linear Associator.

$$6. \mathbf{L}\mathbf{A}\mathbf{y} = \begin{bmatrix} y_1^1 & y_1^2 & \cdots & y_1^p \\ y_2^1 & y_2^2 & \cdots & y_2^p \\ \vdots & \vdots & \cdots & \vdots \\ y_n^1 & y_n^2 & \cdots & y_n^p \end{bmatrix}$$

Recall phase is described through the following algorithm:

1. Present, at the input to Stage 1, a vector from the fundamental set $\mathbf{x}^\mu \in A^n$, for some index $\mu \in \{1, \dots, p\}$.

2. Build vector: $\mathbf{u} = \sum_{i=1}^p \mathbf{h}^i$

3. Do expansion: $\mathbf{F} = \tau^c(\mathbf{x}^\mu, \mathbf{u}) \in A^{n+p}$

4. Obtain vector: $\mathbf{R} = \mathbf{V}\Delta_\beta \mathbf{F} \in A^{n+p}$

5. Do contraction: $\mathbf{r} = \tau^c(\mathbf{R}, n) \in A^p$

If \mathbf{r} is one-hot vector, it is assured that $k = \mu$, then $\mathbf{y}^\mu = \mathbf{L}\mathbf{A}\mathbf{y} \cdot \mathbf{r}$. Stop. Else:

6. For $1 \leq i \leq p$: $w_i = u_i - 1$

7. Do expansion: $\mathbf{G} = \tau^c(\mathbf{x}^\mu, \mathbf{w}) \in A^{n+p}$

8. Obtain a vector: $\mathbf{S} = \mathbf{\Lambda}\nabla_\beta \mathbf{G} \in A^{n+p}$

9. Do contraction: $\mathbf{s} = \tau^c(\mathbf{S}^\mu, n) \in A^p$

10. If \mathbf{s} is zero-hot vector then it is assured that $k = \mu$, $\mathbf{y}^\mu = \mathbf{L}\mathbf{A}\mathbf{y} \cdot \bar{\mathbf{s}}$, where $\bar{\mathbf{s}}$ is the negated vector of \mathbf{s} . Stop. Else:

11.12. Do operation $\mathbf{r} \wedge \bar{\mathbf{s}}$, where \wedge is the symbol of the logical AND operator, so $\mathbf{y}^\mu = \mathbf{L}\mathbf{A}\mathbf{y} \cdot (\mathbf{r} \wedge \bar{\mathbf{s}})$. STOP.

The process in the contrary direction, which is presenting pattern \mathbf{y}^k ($k = 1, \dots, p$) as input to the Alpha-Beta BAM and obtaining its corresponding \mathbf{x}^k , is very similar to the one described above. The task of Stage 3 is to obtain a one-hot vector \mathbf{h}^k given a \mathbf{y}^k . Stage 4 is a modified Linear Associator built in similar fashion to the one in Stage 2.

4 Experiments and Results

The algorithm was implemented with the programming language Microsoft Visual C# 2008 Express Edition® and was tested on a PC with Intel Pentium 4® processor and 1 GB of RAM memory, the operating system was Microsoft Windows XP Professional®.

The information was taken from Oxford Parkinson's Disease Telemonitoring Dataset. This dataset is composed of a range of biomedical voice measurements from 42 people with early-stage Parkinson's disease recruited to a six-month trial of a telemonitoring device for remote symptom progression monitoring.

The description of the attributes is shown in Table 6. The total number of recordings is 195, from which 147 belong to the Parkinson class (class 1) and 48 to the non-Parkinson class (class 2). The database was created in June 2008.

The first step for the classification task is the designing of the Alpha-Beta BAM. Therefore, the memory has to be trained with the set of patterns which are contained in the database. As we can observe in Table 6, the attributes are represented by integer and real numbers, thus they have to be codified in order to obtain patterns with binary values. Each row in the database is binarized as follows (see figure 6): each feature is codified with the Johnson-Möbius code, and then they are concatenated to form a single vector which represents the input pattern x^k , in this case $k = 1, 2, \dots, n$ with $n=195$.

In this work, the algorithm of the learning phase (introduced in Section 3) was just applied for one direction: $x \rightarrow y$, because this application does not have a bidirectional behavior.

Due to Alpha-Beta BAM shows a correct recall for all the training patterns, we first trained the memory with the 195 records.

Table 6. Attribute information

Attribute	Description
subject#	Integer that uniquely identifies each subject
age	Subject age
sex	Subject gender '0' - male, '1' - female
test_time	Time since recruitment into the trial. The integer part is the number of days since recruitment
motor_UPDRS	Clinician's motor UPDRS score, linearly interpolated
total_UPDRS	Clinician's total UPDRS score, linearly interpolated
Jitter:%, Abs, RAP, PPQ5, DDP	Several measures of variation in fundamental frequency
Shimmer: dB, APQ3, APQ5, APQ11, DDA	Several measures of variation in amplitude
NHR,HNR	Two measures of ratio of noise to tonal components in the voice
RPDE	A nonlinear dynamical complexity measure
DFA	Signal fractal scaling exponent
PPE	A nonlinear measure of fundamental frequency variation

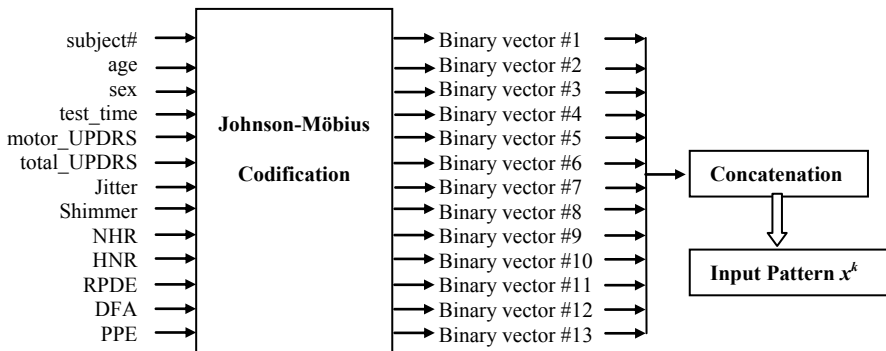


Fig. 6. Binarization of attributes using Johnson-Möbius codification

In the recalling phase (introduced in Section 3), each pattern is codified using the process showed in figure 6 and presented to the Alpha-Beta BAM. As such as we expected every pattern was correctly classified.

For testing the performance of Alpha-Beta BAM algorithm, three methods were used: Hold-Out, K-Fold Cross-Validation and Leave One Out.

Hold-Out Test

The complete database was divided into two sets, one of them was the training set and the other was the testing set. We selected randomly the elements contained in each set. The size of the sets was varying from 2% for training and 98% for testing to 98% for training and 2% for testing. For each size we performed 15 calculations. Table 7 shows the results obtained.

Most of the works performing classification take the effectiveness with the size of the training set of 80%. In our case, we obtained a 98.13 % for that size.

Table 7. Results from Hold-Out test

Size for training (%)	Effectiveness (%)
2	28.73
10	61.53
20	82.40
30	86.20
40	88.27
50	91.27
60	96.00
70	96.87
80	98.13
90	98.93
98	99.87

K-fold Cross-Validation Test

In this case, the whole database was divided into 10 sets (then $K=10$) with the same number of elements. We attempted to maintain the proportion of the number of the records in both classes because the database is not balanced. Therefore, we had 9 sets with 20 records: 15 from Parkinson class and 5 from non-Parkinson class, and 1 set with 15 records: 12 from class 1 and 3 from class 2.

Table 8. Results from K-fold Cross-Validation test with $K=10$

Set	Effectiveness (%)
K1	98
K2	98
K3	97
K4	100
K5	100
K6	96
K7	98
K8	98
K9	99
K10	98
Average	98.2

We took the first set K1 for testing and the remaining sets were for training, afterwards, the set K2 was used for testing and the others for training and so on. The same way, we performed 20 calculations for every test. Table 8 shows the results.

We can observe from the table that the worst effectiveness is equal to 96% while the best was of 100% giving an average of 98.2%.

Leave One Out Test

For this test, the training set contained all the records from the database but one, which was used for testing. Therefore, we performed 195 calculations. With this test, the effectiveness was of 95.19%. Table 9 shows the results.

Table 9. Results from Leave One-Out test

Class	Effectiveness
Parkinson	94.55 (8 mistakes)
Non-Parkinson	95.83 (2 mistakes)
Total	95.19%

Table 10 shows the results from the three testing methods.

Table 10. Results from K-fold Cross-Validation test with K=10

Method	Effectiveness (%)
Hold Out (80% training)	98.13
K-fold Cross-Validation (K=10)	98.2
Leave One Out	95.19
Average	97.17

From Table 10 we can observe that the average of effectiveness of the 3 testing methods was of 97.17%. But we cannot assure this result is good without compare it with other classification algorithms. Therefore, in Table 11 the results from four other methods using the same database are showed. They used different approaches to the one used in this work: Probabilistic Neural Networks (PNN), Mutual Information (MI), Support Vector Machine (SVM) and a combination of Neural Networks (NN) with SVM.

Table 11. The results of effectiveness of the five methods: PNN,MI,SVM, NN-SVM and Alpha-Beta BAM

Method	Effectiveness (%)
Probabilistic Neural Network	81.74
Mutual Information	81.53
Support Vector Machine	62.217
NN-SVM	93.33
Alpha-Beta BAM	97.17

The work that used MI approach tested its algorithm with Leave-One-Out test obtaining 81.53% of classification. With the same testing algorithm we obtained the 95.19% of effectiveness. The result from PNN algorithm was of 81.74% of effectiveness using Hold Out test with 70% for training and 30% for testing. For the same test we obtained 96.87% of effectiveness. In the work where SVM is used, authors applied K-fold Cross-Validation varying K from 2 to 10, and they find the best result (62.217% of effectiveness) with K=10. Alpha-Beta BAM approach shows, with the same test conditions, an effectiveness of 98.2. Finally, the effectiveness of NN-SVM method was of 93.33% but the testing method is not mentioned. However, any of the results from the three testing methods achieved by Alpha-Beta BAM is higher than 93.33%.

5 Conclusions

Parkinson's disease has become a frequent neurodegenerative disease which has affected many people in the world. Therefore, it is important to rely on a system capable of diagnosing patients with Parkinson's disease.

Alpha-Beta Associative models have shown to be an option as a tool for many applications, and recently they have been used specifically as classifiers improving the results by means of Johnson-Möbius modified code.

We implemented Alpha-Beta BAM algorithm together with Johnson-Möbius codification to classify patients with Parkinson's disease. We tested Alpha-Beta BAM using three methods: Hold Out, K-fold Cross-Validation (K=10) and Leave One Out. We compared the obtained results with other different approaches: Probabilistic Neural Networks, Mutual Information, Support Vector Machine and a combination of Neural Networks and SVM. Our approach showed the best results of performance surpassing the Mutual Information approach with 13.66% of effectiveness and Probabilistic Neural Networks with 15.13% of effectiveness. The difference between SVM and our BAM was of 35.983% of accuracy, and Alpha-Beta BAM showed better results than NN-SVM.

These results reassert the fact that Associative Models are a good alternative for the task of classification over other conventional approaches.

Acknowledgments. The authors would like to thank the Instituto Politécnico Nacional (COFAA and SIP), and SNI for their economical support to develop this work.

References

1. Goetz, C.G.: Early Iconography of Parkinson's Disease. In: Handbook of Parkinson's Disease, 4th edn., Informa Healthcare, New York (2007)
2. Factor, S.A., Weiner, W.J.: Parkinson's Disease: Diagnosis and Clinical Management, 2nd edn. Demos, New York (2008)
3. Lieberman, A.: 100 Questions and Answers about Parkinson Disease. Jones and Bartlett Publishers, Sudbury (2003)

4. Jancovic, J.: Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry* 79, 368–376 (2008)
5. Adler, C.H., Ahlskog, J.E.: *Parkinson's Disease and Movements Disorders: Diagnosis and Treatment Guidelines for the Practicing Physician*. Humana Press, New Jersey (2000)
6. Unified Parkinson Disease Rating Scale (UPDRS), <http://www.uninet.edu/neurocon/neurologia/escalas/parkinson.html#UPDRS>
7. World Health Organization, <http://www.who.int>
8. National Parkinson Foundation, <http://www.parkinson.org/parkinson-s-disease.aspx>
9. Acton, P.D., Newberg, A.: Artificial network classifier for the diagnosis of Parkinson's disease using [99mTc]TRODAT-1 and SPECT. *Phys. Med. Biol.* 51(12), 3057–3066 (2006)
10. Ericsson, A., Lonsdale, M.N., Astrom, K., Edenbrandt, L., Friberg, L.: Decision Support System for the Diagnosis of Parkinson's Disease. In: Kalviainen, H., Parkkinen, J., Kaarna, A. (eds.) SCIA 2005. LNCS, vol. 3540, pp. 740–749. Springer, Heidelberg (2005)
11. Yoshida, H., Nakagawa, K., Anai, H., Horimoto, K.: Exact Parameter Determination for Parkinson's Disease Diagnosis with PET Using an Algebraic Approach. In: Anai, H., Horimoto, K., Kutsia, T. (eds.) Ab 2007. LNCS, vol. 4545, pp. 110–124. Springer, Heidelberg (2007)
12. Keijsers, N.L.W., Horstink, M.W.I.M., Gielen, C.C.A.M.: Automatic, unsupervised classification of dyskinesia in patients with Parkinson's Disease. In: Kaynak, O., Alpaydin, E., Oja, E., Xu, L. (eds.) ICANN 2003 and ICONIP 2003. LNCS, vol. 2714, Springer, Heidelberg (2003)
13. Djurić-Jovičić, M., Jovičić, N.S., Milovanović, I., Radovanović, S., Kresojević, N., Popović, M.B.: Classification of walking patterns in Parkinson's disease patients based on inertial sensor data. *Neural Network Applications in Electrical Engineering (NEUREL)*, 3–6 (2010)
14. Parkinsons Telemonitoring Data Set, <http://archive.ics.uci.edu/ml/datasets/Parkinsons>
15. Ene, M.: Neural network-based approach to discriminate healthy people from those with Parkinson's disease. *Annals of the University of Craiova, Math. Comp. Sci. Ser.* 35, 112–116 (2008)
16. Sakar, O., Kursun, O.: Telediagnosis of Parkinson's Disease Using Measurements of Disphonia. *J. Med. Syst.* 34, 591–599 (2010)
17. Bhattacharya, I., Bhatia, M.P.S.: SVM classification to distinguish Parkinson disease patients. In: *Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing A2CWic 2010, India* (2010)
18. Gil, D., Johnson, M.: Diagnosing Parkinson by using Artificial Neural Networks and Support Vector Machines. *Global Journal of Computer Science and Technology* 9(4), 63–71 (2009)
19. Acevedo, E., Yáñez, C., López, I.: Alpha-Beta Bidirectional Associative Memories: Theory and Applications. *Neural Processing Letters* 26, 1–40 (2007)
20. Ritter, G.X., Sussner, P., Diaz de León, J.L.: Morphological Associative Memories. *IEEE Transactions on Neural Networks* 9, 281–293 (1998)
21. Yáñez-Márquez, C.: *Associative Memories Based on Order Relations and Binary Operators* (in Spanish). PhD Thesis. Centro de Investigación en Computación, Mexico (2002)
22. Mano, M.: *Diseño digital*, 16–26, 292–294. Prentice Hall, Englewood Cliffs (2001)
23. Flores, R.: *Johnson-Möbius modified code-based Alpha-Beta Associative Memories* (in Spanish). MD thesis, Centro de Investigación en Computación, Mexico (2006)