ELSEVIER

Information Sciences 142 (2002) 249–273

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

Morphogenic neural networks encode abstract rules by data

G. Resconi [a], A.J. van der Wal [b]

[a] *Department of Mathematics, Catholic University, Via Trieste 17, Brescia, Italy*

[b] *Electrical Engineering Department, C.J. Drebbel Institute for Mechatronics and Control Laboratory, University of Twente, P.O. Box 217, NL-7500 AE Enschede, The Netherlands*

Abstract

The classical McCulloch and Pitts neural unit is widely used today in artificial neural networks (NNs) and essentially acts as a non-linear filter. Classical NN are only capable of approximating a mapping between inputs and outputs in the form of a lookup table or "black box" and the underlying abstract relationships between inputs and outputs remain hidden. Motivated by the need in the study on neural and neurofuzzy architectures, for a more general concept than that of the *neural unit*, or *node*, originally introduced by McCulloch and Pitts, we developed in our previous work the concept of the *morphogenetic neural* (MN) *network*. In this paper we show that in contrast to the classical NN, the MN network can encode abstract, symbolic expressions that characterize the mapping between inputs and outputs, and thus show the internal structure hidden in the data. Because of the more general nature of the MN, the MN networks are capable of abstraction, data reduction and discovering, often implicit, relationships. Uncertainty can be expressed by a combination of evidence theory, concepts of quantum mechanics and a morphogenetic neural network. With the proposed morphogenetic neural network it is possible to perform both rigorous and approximate computations (i.e. including semantic uncertainty). The internal structure in data can be discovered by identifying "invariants", i.e. by finding (generally implicit) dependencies between variables and parameters in the model. © 2002 Elsevier Science Inc. All rights reserved.

Keywords: Morphogenetic neuron; Neural network; Quantum computing; Filtering; Evidence theory; Synergy; Orthogonality; Invariant; Internal structure of data; Semantic uncertainty

E-mail addresses: resconi@numerica.it (G. Resconi), a.j.vanderwal@el.utwente.nl (A.J. van der Wal).

PII: S0020-0255(02)00168-8

1. Introduction

Recent research indicates the need for the introduction of a more general concept than that of the *neural unit*, or *node*, dating back to the pioneering work by McCulloch and Pitts [1]. In this paper we use neural units that act as input–output filters of a much more general form, more suitable for modelling active dendrites (cf. e.g. [2–4]). In Section 2 we will first shortly recapitulate the concept of the morphogenetic neuron (MN) introduced in our earlier work. Next, in Section 3 an overview is given of quantum computing to illustrate the parallels between neural computation and quantum computation, as illustrated by the concept of linear superposition. The connection between neural networks (NNs) and quantum computing is made in Section 4. In Section 5. the morphogenetic filter is introduced in orthogonal and non-orthogonal basis functions; and the role of the scalar product in both the MN and in biological neural systems is discussed. Finally the existence of invariants within the context is defined by the fundamental tensor. In Section 6 the algorithm to generate abstract rules is formulated. In Section 7, we discuss tensor invariants and in Sections 8 and 9 we give examples of the use of morphogenetic networks as filters to extract symbolic expressions from numerical data in robotics which are given. It is shown that the morphogenetic neural network is capable of discovering internal structure in measured numerical data. Finally, in Section 10 we conclude by discussing how semantic uncertainty (evidence theory and quantum indeterminism) can be included in the theory of the morphogenetic neuron. In Section 11 we extend the scalar product to include the cognitive aspects in evidence theory.

2. The morphogenetic neuron

In [5–8] we introduced a generalization of the concept of neural unit, which has been named *morphogenetic neuron*. From an abstract point of view, the latter is a generic analog input–output device through which two elementary operations are possible: the "Write" and "Read" operation, i.e.

(1) the operation ("Write"), starting from a suitable *reference space* and from the *constraints* to be satisfied generates the weights at the synapse,

(2) the operation computation ("Read"), starting from the weights to construct a suitable reply that satisfies the imposed constraints.

The attribute "morphogenetic" was chosen because, according to our view, the existence of a set of constraints on a reference space, spanned, e.g. by all possible values of a number of input or state variables, induces the "shape" of the morphogenetic neuron within this space.

The name "neuron" was adopted because the activation function of such a device is characterized, in the same way as usual neural units, by a *bias* *potential* and by a *weighted sum* of suitable (in general non-linear) functions of morphogenetic fields. The read function is a linear superposition, with the weights obtained in the write operation, of elementary functions or states. We show that simple algorithms exist to compute the weights to implement three fundamental operations. The implementation of a given input–output transfer function, the operation of filter decomposition of data in given prototype data with different weights, and the generation of instructions that transform data (transformation of the space context). With these three types of operations we can construct any possible computation.

The operation of the morphogenetic neuron can generally be described as a two-step process. First the neuron learns its structural context, i.e. the *n*-dimensional vector space of functions. Next it learns how to encode abstract rules as scalar invariant expressions within this context.

In the following we establish an interesting parallelism between the quantum computation and the neuron computation, based on the observation that also the quantum computer is based on the weighted superposition.

3. Quantum computing

Quantum computation is the extension of classical computation to the processing of quantum information, using quantum systems such as individual molecules, atoms, or elementary particles such as photons. For an introduction, see e.g. [9,10]. It has the potential to bring about a spectacular revolution in computer science. Current-day electronic computers are not fundamentally different from purely mechanical computers: the operation of such a Turing machine can be described completely in terms of classical physics. By contrast, computers could in principle be built to benefit from genuine quantum phenomena that have no classical analogon such as *entanglement and interference*, sometimes providing exponential speed-up compared with classical computers [11].

All computers manipulate information, and the unit of quantum information is the quantum bit, or *qubit*.

Classical bits can take either the values 0 or 1, but qubits can form a linear *superposition* of the two classical states. If we denote the classical bits by $|0\rangle = \psi_0$ and $|1\rangle = \psi_1$ where $P_0(x,y) = |\psi_0(x,y)|^2$, like in quantum mechanics, is the probability that the qubit in the state zero should be in the position x,y.

When the state changes from zero to one, the probability that the qubit should be in the state one and in the position x,y is $P_1(x,y) = |\psi_1(x,y)|^2$. A quantum bit can be in *any* state

$$\alpha|0\rangle + \beta|1\rangle,$$

where α and β are complex numbers called *amplitudes* subject to the normalization condition

$$|\alpha|^2 + |\beta|^2 = 1.$$

Any attempt to measure qubits induces an irreversible disturbance. For example, the most direct measurement on $\alpha|0\rangle + \beta|1\rangle$ results in the qubit making a probabilistic decision: with probability $|\alpha|^2$ the result of the measurement becomes $|0\rangle$ and with complementary probability $|\beta|^2$ the outcome becomes $|1\rangle$; in either case the measurement apparatus tells us which choice has been taken, but all previous information of the original amplitudes α and β is lost and the outcome of the computation is probabilistic, in contrast to the classical, deterministic, sequential operation of a Turing machine.

Unlike classical bits, where a single string of length n, consisting of zeros and ones suffices to completely describe the state of n bits, a physical system of n qubits requires 2^n complex numbers to describe its state. For example, a two-qubit state can be represented as:

$$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

for arbitrary complex numbers α, β, γ, and δ, subjected only to the normalization constraint

$$|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1.$$

Another feature of qubits is the property of *entanglement.* Some special states such as

$$(|01\rangle - |10\rangle)/\sqrt{2}$$

cannot be factored.

When these two qubits are measured, they yield either 0 and 1 or 1 and 0, with equal probabilities 1/2 but which of these two outcomes will occur is not determined until the measurement is actually performed. This has no classical analogous. Computers that thrive on entangled quantum information could run exponentially faster than classical computers because n qubits require 2^n numbers for their description. A few simple logic operations by unitary transformations on these qubits can affect all 2^n numbers through the use of quantum parallelism and quantum interference. Examples of this are e.g. Shor's algorithm [11] and Grover's search algorithm for databases.

4. Neural network and quantum computation

A model of a classical NN is given in Fig. 1.

When the unit i receives input from the units j the activation of the unit i, u_i, is given by

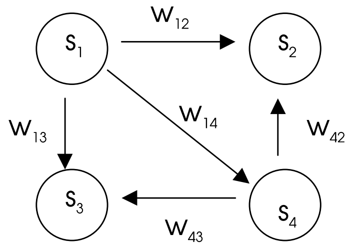

Fig. 1. NN, where s_i represents the output (state) of neuron i and the w_{ij} are the connecting weights.

$$u_i = \sum_j w_{i,j} s_j, \qquad j = 1, 2, \ldots, n, \tag{1}$$

i.e. the weighted sum of the inputs s_j. The activation u_i is subsequently used to determine the output (i.e. the new state) of the neuron i by evaluating a non-linear threshold function f

$$s_i^{\text{new}} = f(u_i).$$

In a net of real neurons the $w_{i,j}$ represent the weights of the synapses. We note that the inputs s_j are independent from each other and at any synapse j only one input is active. Recent experiments [20–22] show that this simple model of a neural net is not realistic. At every synapse many different inputs interact in a synergistic way in order to activate an axon. In order to model the real single neuron better we propose a different model of a neural net, *the morphogenetic model*, schematically given in Fig. 2. In Eq. (2) we give the formal *definition of the morphogenetic neuron*.

The activation u_i of the morphogenetic neuron equals the output and is given by

$$u_i = \sum_j w_{i,j} \psi_j(s_1, \ldots, s_n), \quad j = 1, 2, 3, \ldots, 2^n. \tag{2}$$

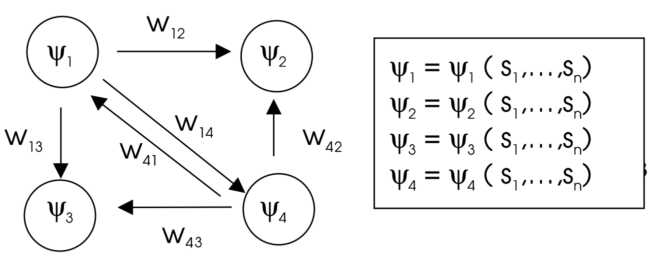

Fig. 2. Morphogenetic neural network built of states ψ_i connected by weights w_{ij}.

The parameters $w_{i,j}$ are the synaptic weights and ψ_j represents the state of the neuron j [21,22]. We remark that this MN is more general than the classical neuron, since the non-linearity of the classical neuron can be modelled as a superposition of non-linear functions ψ, e.g. a Heaviside function could be expanded in sine and cosine functions via the Fourier transform. In the classical neuron the superposition is applied before the non-linear function, whereas in the MN we superimpose non-linear functions.

As an example we consider a neuron with two inputs s_1 and s_2

$$u = w_1\psi_1(s_1, s_2) + w_2\psi_2(s_1, s_2) + w_3\psi_3(s_1, s_2) + w_4\psi_4(s_1, s_2),$$

where $\psi_1(s_1, s_2)$, $\psi_2(s_1, s_2)$, $\psi_3(s_1, s_2)$, $\psi_4(s_1, s_2)$ are the elementary states of the neuron. The elementary states of the neuron can be written also in the following way: $\psi_{00}(s_1, s_2)$, $\psi_{10}(s_1, s_2)$, $\psi_{01}(s_1, s_2)$, $\psi_{11}(s_1, s_2)$, so that the morphogenetic neuron model can be written as

$$u = w_1\psi_{00}(s_1, s_2) + w_2\psi_{01}(s_1, s_2) + w_3\psi_{10}(s_1, s_2) + w_4\psi_{11}(s_1, s_2).$$

We remark that superposition of the morphogenetic neuron states has an analogy with the superposition of states in quantum mechanics. To stress the connection between the quantum mechanics and the morphogenetic neuron model we rewrite Eq. (2) using the Dirac formalism in the following way:

$$\begin{aligned} u_i &= \sum_j w_{i,j}|j_1, j_2, \ldots, j_n\rangle \\ &= w_{i,1}|0, 0, \ldots, 0\rangle + w_{i,2}|1, 0, \ldots, 0\rangle + \cdots + w_{i,2^n}|1, 1, \ldots, 1\rangle, \end{aligned} \tag{3}$$

where we used the Dirac notation from quantum mechanics

$$\psi_j(s_1, \ldots, s_n) = |j_1, j_2, \ldots, j_n\rangle. \tag{4}$$

Example 1. If the connection terms $w_{i,j}$ form a unitary matrix, Eq. (3) represents a unitary transformation in quantum mechanics and can be interpreted as an instruction in the quantum computer

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{bmatrix} \begin{bmatrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{bmatrix} = \begin{bmatrix} a_{1,1}|00\rangle + a_{1,2}|01\rangle + a_{1,3}|10\rangle + a_{1,4}|11\rangle \\ a_{2,1}|00\rangle + a_{2,2}|01\rangle + a_{2,3}|10\rangle + a_{2,4}|11\rangle \\ a_{3,1}|00\rangle + a_{3,2}|01\rangle + a_{3,3}|10\rangle + a_{3,4}|11\rangle \\ a_{4,1}|00\rangle + a_{4,2}|01\rangle + a_{4,3}|10\rangle + a_{4,4}|11\rangle \end{bmatrix},$$

when

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

we obtain the transformation

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{bmatrix} = \begin{bmatrix} 1|00\rangle + 0|01\rangle + 0|10\rangle + 0|11\rangle \\ 0|00\rangle + 1|01\rangle + 0|10\rangle + 0|11\rangle \\ 0|00\rangle + 0|01\rangle + 0|10\rangle + 1|11\rangle \\ 0|00\rangle + 0|01\rangle + 1|10\rangle + 0|11\rangle \end{bmatrix} = \begin{bmatrix} |00\rangle \\ |01\rangle \\ |11\rangle \\ |10\rangle \end{bmatrix}.$$

That realize by the superposition the two Boolean functions

X	Y	Boolean function X
0	0	0
0	1	0
1	0	1
1	1	1

X	Y	XOR
0	0	0
0	1	1
1	0	1
1	1	0

The unitary transformations used in this example affect qubits $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ in a synchronic way by the use of quantum parallelism and quantum interference.

This unitary transformation can formally be represented by the graph depicted in Fig. 3.

We know that in quantum mechanics the superposition of the states gives a new coherent state in which the previous states are mixed. In the morphogenetic neuron the superposition of synchronous inputs creates the output state in which the previous input states are mixed.

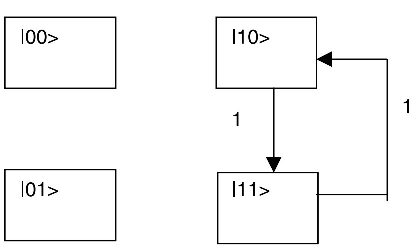

Fig. 3. Quantum transformation represented as a network.

5. Filtering with the morphogenetic neuron

5.1. Orthogonality

Given the elementary states we can generate a mixed state when we know the coefficients, but from the mixed state is difficult to obtain the coefficients of the elementary states (tuning process), e.g. by the superposition of the states $\psi_1,\ \psi_2,\ \psi_3,\ \psi_4$ we may generate the function $\psi = \psi_1 - 2\psi_2 + \psi_3 + 1/2\psi_4$. The elementary states ψ_i $(i = 1, \dots, 4)$, generate a basis for an n-dimensional Hilbert space, where $n = 4$ is the number of basis states. We know that the scalar product for continuous basis functions is defined as:

$$\psi_i(x) \cdot \psi_j(x) = \int_\Omega \psi_i(x)\psi_j(x)\,\mathrm{d}x.$$

In numerical calculations the continuous basis is often replaced by a discrete one, and the scalar product is then approximated by

$$\psi_i(x) \cdot \psi_j(x) \approx \sum_k \psi_i(x_k)\psi_j(x_k).$$

We note that in case a discrete basis exists, this approximation of the scalar product reduces to an exact definition. When the basis exists in the ordinary Hilbert space the states are orthogonal when the metric tensor $G = \{g_{i,k}\}$ satisfies

$$g_{i,k} = G = \begin{bmatrix} \psi_1 \cdot \psi_1 & \psi_1 \cdot \psi_2 & \cdots & \psi_1 \cdot \psi_n \\ \psi_2 \cdot \psi_1 & \psi_2 \cdot \psi_2 & \cdots & \psi_2 \cdot \psi_n \\ \cdots & \cdots & \cdots & \cdots \\ \psi_n \cdot \psi_1 & \psi_n \cdot \psi_2 & \cdots & \psi_n \cdot \psi_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$
$$= \delta_{j,k}.$$

We know that for a unitary transformation S we have $SS^{\mathrm{T}} = 1$, where S^{T} is the transpose of S.

For the unitary transformation S we then have

$$\psi_k \cdot \psi_j = S\psi_k \cdot S\psi_j = SS^{\mathrm{T}}\psi_k \cdot \psi_j.$$

It follows that the scalar product is invariant under S. This is an example of a tensor invariant (see Section 7). When $G = 1$ and when the basis set of states is orthogonal, S transforms the basis set of states to another basis set that is again orthogonal. The metric $g_{i,k}$ does not change. We conclude that the states $\psi_1, \psi_2, \dots, \psi_n$ generate an "entity" or space that we cannot separate in parts and that when the states are orthogonal no relation exists between the states.

When the basis states are orthogonal it is possible to filter (decompose) a mixed state in its elementary states by using the scalar product. We can see this as follows.

Let ψ be a mixed state given by

$$\psi = \sum_j \alpha_j \psi_j.$$

We can obtain the projections along the basis functions ψ_k by taking the scalar product with ψ_k

$$\psi \cdot \psi_k = \psi_k \cdot \sum_j \alpha_j \psi_j = \sum_j \alpha_j \psi_j \cdot \psi_k = \sum_j \alpha_j \delta_{j,k} = \alpha_k.$$

We note that in the mixed state we can have a large number of basis states ("features") superposed with different intensities. A mixed state thus compresses a huge amount of information in one function. Given the basis state ψ_k we can select inside the mixed state the intensity of the state ψ_k. The function ψ_k acts as a feature selector, since the scalar product with ψ gives the feature strength α_k

$$\alpha_k = \psi \cdot \psi_k.$$

Finally we note that the computation of the α_k has low computational cost. In contrast, recording all the states and coefficients and finding the desired coefficient α_k amongst all the others are computationally expensive processes. Next we will investigate how the orthogonal basis functions ("features") behave under a transformation R. If the transformation R changes the basis function ψ_k to $\psi_k^* = \sum_j R_{k,j} \psi_j$ we obtain:

$$\psi^* = \sum_k \alpha_k^* \psi_k^* = \sum_k \alpha_k^* \sum_j R_{k,j} \psi_j = \sum_j \left[\sum_k R_{k,j} \alpha_k^* \right] \psi_j = \sum_j \alpha_j \psi_j,$$

$$\sum_k R_{k,j} \alpha_k^* = \alpha_j.$$

Example 2. Given the operator R

$$R\psi = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} \psi_1 \cos(\theta) - \psi_2 \sin(\theta) \\ \psi_1 \sin(\theta) + \psi_2 \cos(\theta) \end{bmatrix} = \begin{bmatrix} \psi_1^* \\ \psi_2^* \end{bmatrix}$$

with the scalar product and orthogonal states ψ_1 and ψ_2 we have

$$\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.$$

5.2. Neural pattern recognition and scalar product in biological systems

Research results from neurobiology and neurophysiology suggest that the concept of morphogenetic neuron can be useful to describe what happens in human or animal brains when doing pattern recognition tasks. Studies on auditory cortex of bats indicated, e.g. the presence of neurons sensitive to the frequency distribution of the input signal (cf. [12]). Moreover, the cortex output can be viewed more as an output function rather than as a sum of single neural impulses. The study of the role of hippocampal cells as determinants of the spatial memory of a rat and from the analysis of motion detectors of insects further corroborate these findings [13,14]. In this section we will describe a possible application of morphogenetic neurons in the design of an artificial device able to classify input patterns. Such a device is based on models of pattern recognition proposed within Cognitive Psychology (see e.g. [4]). One of these models is based on the so-called "template matching" process. According to such a model every recognition system contains, hard-wired into it, a number of "template" patterns. A given input pattern is compared, through a matching operation, with each stored template. The template corresponding to the best match specifies automatically the class to which the input pattern belongs.

The scalar product defined previously gives us the opportunity to connect additive superposition of functions with the scalar product. We know that in the biological neuron network the scalar product

$$P_n = \mathbf{V}_1 \cdot \mathbf{V}_n, \quad n = 1, 2, 3, \ldots, N,$$

between the external stimulus vector $\mathbf{V}_1$ and n internal vectors or reference vectors is the basic instrument for a tuning process by which we can select a desired feature inside the stimulus. Fig. 4 illustrates the scalar product as

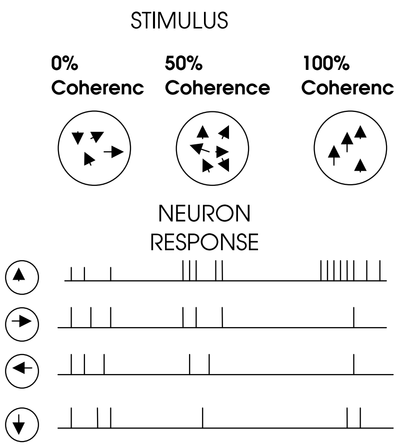

Fig. 4. The measured response of a biological network of neurons depends on the scalar product of the external stimulus and internal reference prototypes (templates).

mechanism for coupling stimuli with pre-programmed templates in order to generate a neuron response.

5.3. Non-orthogonal basis set

When two basis states ψ_1 and ψ_2 are *not* orthogonal, their scalar product does not vanish

$$\psi_1 \cdot \psi_2 = |\psi_1||\psi_2|\cos(\theta), \quad \text{where } |\psi| = \sqrt{\psi \cdot \psi} \text{ and } \theta,$$

is the angle between the two states. When $|\psi_1| = 1$ and $|\psi_2| = 1$ we have

$$\psi_1 \cdot \psi_2 = \cos(\theta).$$

Therefore we conclude that when the elementary states are not orthogonal, i.e. $g_{j,k} \neq \delta_{j,k}$, we cannot use the scalar product to decompose a mixed state into basis states.

This can be illustrated by projecting the mixed state $\psi_1 - 2\psi_2$ onto ψ_2

$$[\psi_1 - 2\psi_2] \cdot \psi_2 = \psi_1 \cdot \psi_2 - 2\psi_2 \cdot \psi_2 = \cos(\theta) - 2 \neq -2.$$

We note that for a mixed state such as $(\alpha_1\psi_1 + \alpha_2\psi_2) = \psi$ the scalar product

$$\begin{aligned} \psi_i \cdot \psi &= \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \cdot (\alpha_1\psi_1 + \alpha_2\psi_2) = \begin{bmatrix} \alpha_1\psi_1 \cdot \psi_1 + \alpha_2\psi_2 \cdot \psi_1 \\ \alpha_1\psi_1 \cdot \psi_2 + \alpha_2\psi_2 \cdot \psi_2 \end{bmatrix} \\ &= \begin{bmatrix} \psi_1 \cdot \psi_1 & \psi_1 \cdot \psi_2 \\ \psi_2 \cdot \psi_1 & \psi_2 \cdot \psi_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \textstyle\sum_j g_{i,j}\alpha_j = G\alpha. \end{aligned}$$

In this case the scalar product does *not* give the coefficients. In order to keep the nice properties found previously for the orthogonal case, we have to *modify* the scalar product in a suitable way. Therefore we suggest the following method: create a new basis consisting of the scalar products of the original basis states with the morphogenetic network given in Fig. 5.

In this way we have the new scalar product

$$P_{j,k} = \sum_i w_{j,i}\psi_i \cdot \psi_k. \tag{5}$$

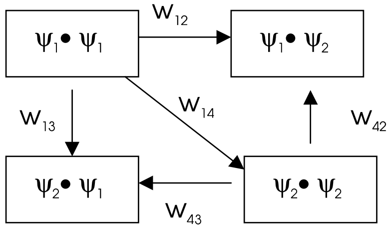

Fig. 5. Morphogenetic neural network used for orthogonalization of basis functions ψ_k.

That can be written in the following way:

$$P = WG = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \begin{bmatrix} \psi_1 \cdot \psi_1 & \psi_1 \cdot \psi_2 & \cdots & \psi_1 \cdot \psi_n \\ \psi_2 \cdot \psi_1 & \psi_2 \cdot \psi_2 & \cdots & \psi_2 \cdot \psi_n \\ \cdots & \cdots & \cdots & \cdots \\ \psi_n \cdot \psi_1 & \psi_n \cdot \psi_2 & \cdots & \psi_n \cdot \psi_n \end{bmatrix}.$$

We now choose the weight matrix W in such a way that P is diagonal

$$\begin{aligned} WG &= \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \begin{bmatrix} \psi_1 \cdot \psi_1 & \psi_1 \cdot \psi_2 & \cdots & \psi_1 \cdot \psi_n \\ \psi_2 \cdot \psi_1 & \psi_2 \cdot \psi_2 & \cdots & \psi_2 \cdot \psi_n \\ \cdots & \cdots & \cdots & \cdots \\ \psi_n \cdot \psi_1 & \psi_n \cdot \psi_2 & \cdots & \psi_n \cdot \psi_n \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{6}$$

it follows that $W = G^{-1}$, where the matrix G is the metric of the state space. The matrix W changes the metric of the space and compensates the errors that we introduce when the states are not orthogonal.

Remark. When the basic functions are not orthogonal, we can write the non-orthogonal set ψ_i as the transformation by R of the orthogonal set of functions η_i. So we have

$$\psi_i = R\eta_i.$$

In this case we have that

$$G = RR^{\mathrm{T}}.$$

For two non-orthogonal, normal states ψ_1 and ψ_2 making an angle θ with each other the metric G tensor is

$$G = \begin{bmatrix} 1 & \cos(\theta) \\ \cos(\theta) & 1 \end{bmatrix}$$

and the compensation matrix W is

$$W = G^{-1} = \begin{bmatrix} 1 & -\cos(\theta) \\ -\cos(\theta) & 1 \end{bmatrix} \frac{1}{1 - \cos^2(\theta)}.$$

Given the mixed state $\psi = \alpha_1\psi_1 + \alpha_2\psi_2$, we have

$$\begin{aligned} W\psi_i \cdot \psi &= W\psi_i \cdot (\alpha_1\psi_1 + \alpha_2\psi_2) \\ &= \frac{1}{1-\cos^2(\theta)} \begin{bmatrix} 1 & -\cos(\theta) \\ -\cos(\theta) & 1 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} \cdot (\alpha_1\psi_1 + \alpha_2\psi_2) \\ &= \frac{1}{1-\cos^2(\theta)} \begin{bmatrix} 1 & -\cos(\theta) \\ -\cos(\theta) & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 + \alpha_2\cos(\theta) \\ \alpha_1\cos(\theta) + \alpha_2 \end{bmatrix} \\ &= \frac{1}{1-\cos^2(\theta)} \begin{bmatrix} 1 & -\cos(\theta) \\ -\cos(\theta) & 1 \end{bmatrix} \begin{bmatrix} 1 & \cos(\theta) \\ \cos(\theta) & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \end{aligned}$$

and we arrive at the classical scalar products between orthonormal states ψ_j and ψ_k.

We conclude that for the non-orthogonal set of states the scalar product $\psi_i \cdot \psi$ can be written as

$$\psi_i \cdot \psi = \psi_i \cdot \sum_j \alpha_j\psi_j = \sum_j \alpha_j\psi_i \cdot \psi_j = \sum_j G_{i,j}\alpha_j.$$

So

$$\sum_i G^{-1}_{k,i} \sum_j G_{i,j}\alpha_j = G^{-1}G\alpha = \alpha$$

and we obtain the coefficients α as in the orthogonal set of functions.

When $W = RG^{-1}$ we have

$$RG^{-1}G\alpha = R\alpha.$$

When R is a unitary transformation $RR^{\mathrm{T}} = 1$, we move from one orthogonal to another orthogonal set of functions. When R is not unitary, we move from an orthogonal set of functions, to a non-orthogonal set of functions.

5.4. Dependency of states: invariant relations

When the determinant of G is zero, the inverse matrix G^{-1} does not exist. In this case the states are *linear dependent*

$$\sum \beta_h\psi_h = 0, \quad \text{where not all } \beta_k = 0.$$

Let us illustrate this with an example. Suppose that inside the space spanned by $\psi_1,\ \psi_2,\ \psi_3,\ \psi_4$, the internal relation, Eq. (7), holds

$$4\psi_1 + \psi_2 - \psi_3 = 0. \tag{7}$$

One of the major applications of the morphogenetic neuron is to find the internal relations ("invariants") between given numerical data. Given G we must find the minor of G with non-zero determinant. In the previous example when we take ψ_3 as a mixed state of $\psi_1,\ \psi_2$

$$\psi_3 = 4\psi_1 + \psi_2,$$

we move from the space $\psi_1,\ \psi_2,\ \psi_3$ to the subspace $\psi_1,\ \psi_2$. This can be done by filtering the mixed state ψ_3 via the non-orthogonal projection. In this way we determine the internal relation (Eq. (7)). This has been discussed in Section 5.3.

6. The morphogenetic neural network to encode abstract rules.

In this section we will show that the morphogenetic neuron can simulate the well-known property of the brain to extract rules from patterns of data. How human learners extract rules from pattern of data is one of the foci in the investigations of human learning. Here we will focus on the question how we can extract symbolic expressions from raw numerical data. In the following we describe the algorithm to generate invariants from numerical data.

To obtain the invariant forms we constructed the following algorithm:

(1) $G \equiv \Xi^{\mathrm{T}}\Xi$,

where Ξ is the $p \times n$ matrix in Table 1.

And identifying $G_{i,k} = \sum_k \psi_i(x_k)\psi_j(x_k) \approx g_{i,k} = \psi_i \cdot \psi_k = \int_\Omega \psi_i(x)\psi_j(x)\,\mathrm{d}x$.

(2) $W \equiv G^{-1}$.

Given the mixed function $\psi = \sum_j \alpha_j\psi_j$ with the numerical table ψ. In Table 2 we show the sampled data set of the function ψ.

(3) $\mathbf{C} \equiv W\Xi^{\mathrm{T}}\Psi$, where $\mathbf{C}$ equals the vector of the coefficients α_j and $\Xi^{\mathrm{T}}\Psi = \sum_k \psi_i(x_k)\psi(x_k) \approx \psi_i \cdot \psi$.

7. Invariants of a tensor

From tensor algebra (see e.g. [15]) it is known that from a given tensor one can derive infinite many invariant forms. In order to illustrate the concept of

Table 1
Matrix of experimental data in terms of features and samples

	Features			
	ψ_1	ψ_2	$\cdots$	ψ_n
Values sample 1	$\psi_1(x_1)$	$\psi_2(x_1)$		$\psi_n(x_1)$
Values sample 2	$\psi_1(x_2)$	$\psi_2(x_2)$		$\psi_n(x_2)$
Values sample 3	$\psi_1(x_3)$	$\psi_2(x_3)$		$\psi_n(x_3)$
$\vdots$				
Values sample p	$\psi_1(x_p)$	$\psi_2(x_p)$		$\psi_n(x_p)$

Table 2
Vector of samples of general mixed function ψ

	ψ
Values sample 1	$\psi(x_1)$
Values sample 2	$\psi(x_2)$
Values sample 3	$\psi(x_3)$
$\vdots$	
Values sample p	$\psi(x_p)$

invariant, we consider the vector **A**, which has components given by A_i in a rectangular coordinate system K. Let A'_i be the components of **A** in another rectangular coordinate system K'. It is easy to prove by explicit calculation that the scalar product of **A** with itself $\mathbf{A} \cdot \mathbf{A}$, which is the square of the length of **A**, is the same in both coordinate systems

$$A_iA_i = A_1^2 + A_2^2 + A_3^2 = A_1'^2 + A_2'^2 + A_3'^2 = A'_iA'_i.$$

This simple example shows that expressions exist that do not change under transformation from one coordinate system to the other. Such an expression is called an *invariant* of **A**. A vector is a tensor of first order. Tensors of higher order also have invariants. This can be understood by considering the characteristic equation of a second-order tensor g_{ik}

$$\begin{vmatrix} g_{11} - \lambda & g_{12} & g_{13} \\ g_{21} & g_{22} - \lambda & g_{23} \\ g_{31} & g_{32} & g_{33} - \lambda \end{vmatrix} = 0.$$

After expanding this determinant we obtain the equation

$$\lambda^3 - \lambda^2(g_{11} + g_{22} + g_{33}) + \lambda\left(\begin{vmatrix} g_{22} & g_{32} \\ g_{23} & g_{33} \end{vmatrix} + \begin{vmatrix} g_{11} & g_{21} \\ g_{12} & g_{22} \end{vmatrix} + \begin{vmatrix} g_{11} & g_{31} \\ g_{13} & g_{33} \end{vmatrix}\right)$$
$$- \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} = 0.$$

The numbers λ, λ^2, λ^3 are scalars and therefore independent of the choice of coordinate system and so are their coefficients in the characteristic equation; they form invariants I_j of $g_{i,k}$

$$I_1 = g_{11} + g_{22} + g_{33}.$$

I_1 is known as the trace of $g_{i,k}$

$$I_2 = \begin{vmatrix} g_{22} & g_{32} \\ g_{23} & g_{33} \end{vmatrix} + \begin{vmatrix} g_{11} & g_{21} \\ g_{12} & g_{22} \end{vmatrix} + \begin{vmatrix} g_{11} & g_{31} \\ g_{13} & g_{33} \end{vmatrix}, \quad I_3 = \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix}.$$

From these invariants we may derive infinitely many other invariants, e.g.

$$I_1^2 = \left(\sum_{i=1}^{3} g_{ii} \right)^2 \quad \text{or} \quad I_1^2 - 2I_2 = \sum_{i=1}^{3} \sum_{k=1}^{3} g_{ik} g_{ik}.$$

Invariants thus encapsulate part of the internal structure of the transformation represented by the tensor g_{ik} in a form that is not dependent on a specific representation (e.g. a coordinate system). This makes invariants excellent vehicles to extract information reflecting the internal structure of a system (e.g. symbolic expressions). In the following section we will sketch applications to the robotic of the tensor invariants.

8. Application to a SCARA robot

It is easy to write the equations linking the angular position of the links $\mathbf{q} = (\alpha, \beta)$ with the position of the gripper in the Cartesian space $\mathbf{u} = (X, Y)$ (see Fig. 6)

$$X = L1 \cos(\alpha) + L2 \cos(\alpha + \beta),$$
$$Y = L1 \sin(\alpha) + L2 \sin(\alpha + \beta)$$

with the basic functions

$$\psi_1 = L1 \cos(\alpha) + L2 \cos(\alpha + \beta), \quad \psi_2 = L1 \sin(\alpha) + L2 \sin(\alpha + \beta).$$

For $L1 = 1$ and $L2 = 1$, we have the fundamental tensor for the variation of the angle α

$$g_{1,1} = \psi_1 \cdot \psi_1 = \int_0^p (\cos(\alpha) + \cos(\alpha + \beta))^2 \, \mathrm{d}\alpha,$$
$$g_{2,2} = \psi_2 \cdot \psi_2 = \int_0^p [\sin(\alpha) + \sin(\alpha + \beta)]^2 \, \mathrm{d}\alpha,$$

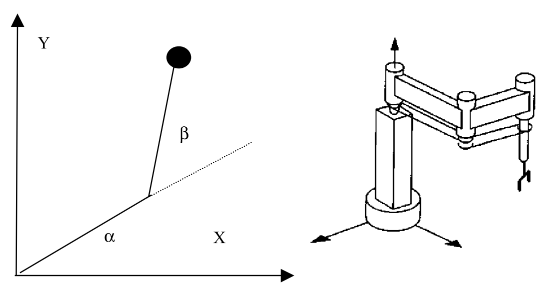

Fig. 6. Scheme of SCARA robot.

evaluating the fundamental tensor and by simple computations we have

$$g_{11} + g_{22} = 2\alpha + 2\alpha \cos \beta. \tag{8}$$

By derivation of Eq. (8) with respect to α we obtain the invariant

$$X^2 + Y^2 = 2(1 + \cos(\beta)).$$

For $L1$ and $L2$ as free parameters we have:

$$X^2 + Y^2 = L1^2 + L2^2 + 2L1L2\cos(\beta).$$

By varying the angle β we obtain the invariant

$$X^2 + Y^2 = L2^2 - L1^2 + 2L1[X\cos(\alpha) + Y\sin(\alpha)].$$

From these two invariants, we can calculate the angles α and β when we know X, Y, $L1$, and $L2$.

9. Application to a spherical robot

Kinematics equations:

$$X = [L1\cos(\alpha) + L2\cos(\alpha + \beta)]\cos(\gamma),$$
$$Y = [L1\cos(\alpha) + L2\cos(\alpha + \beta)]\sin(\gamma),$$
$$Z = [L1\sin(\alpha) + L2\sin(\alpha + \beta)]$$

with the basic functions:

$$\psi_1 = [L1\cos(\alpha) + L2\cos(\alpha + \beta)]\cos(\gamma),$$
$$\psi_2 = [L1\cos(\alpha) + L2\cos(\alpha + \beta)]\sin(\gamma),$$
$$\psi_3 = [L1\sin(\alpha) + L2\sin(\alpha + \beta)].$$

For the variation in γ we have

$$g_{11} = \int_0^\gamma \psi_1(\alpha, \beta, \gamma)^2 \,\mathrm{d}\gamma, \quad g_{22} = \int_0^\gamma \psi_2(\alpha, \beta, \gamma)^2 \,\mathrm{d}\gamma,$$

and obtain the invariant for the angle γ

$$X^2 + Y^2 = [L1\cos(\alpha) + L2\cos(\alpha + \beta)]^2.$$

Let $B = \pm\sqrt{(X^2 + Y^2)}$. We may then reduce the spherical robot to the SCARA robot with the equations:

$$B = L1\cos(\alpha) + L2\cos(\alpha + \beta),$$
$$Z = L1\sin(\alpha) + L2\sin(\alpha + \beta).$$

This set of equations we can solve. By introducing the new variable B, we can reduce the number of variables from three to two. The two variables are the angles α and β of the SCARA robot.

10. The representation of semantic uncertainty in the morphogenetic neuron

Given the proposition "the element a belongs to A" this proposition can be true or false. In the classical information theory the two logical statement (true and false) are represented by a bit whose value is one for "true" and zero for "false". For the proposition "a belongs to A" where is the meaning of the qubit? In fuzzy set theory for the proposition "the element a belongs to A", we associate a fuzzy set where for any element "a" we define a measure of degree of truth $\mu_A(a)$.

To give a meaning to the qubit $\psi_1 = |1\rangle$, we associate the qubit with the basic probability assignment from the evidence theory. When the universal set U equals {0,1}, we have the power set $\mathscr{P} = \{\emptyset, \{0\}\{1\}\{0,1\}\}$. To each element of the power set we associate a basic probability assignment in the evidence theory. The representation of the quantum uncertainty by evidence theory was independently proposed in 1995 by van der Wal [16] and in 1999 by Resconi et al. [17]

$$m(\emptyset) = 0, \quad m(\{0\}) = \langle 0|0\rangle = \int_{\Omega} \psi_0(x)\bar{\psi}_0(x)\,\mathrm{d}x,$$

$$m(\{1\}) = \langle 1|1\rangle = \int_{\Omega} \psi_1(x)\bar{\psi}_1(x)\,\mathrm{d}x,$$

$$m(\{0,1\}) = \langle 01|01\rangle = \int_{\Omega} \psi(x)\bar{\psi}(x)\,\mathrm{d}x$$

with

$$m(\emptyset) + m(\{0\}) + m(\{1\}) + m(\{0,1\}) = 1,$$

where

$$\psi = \alpha|0\rangle + \beta|1\rangle.$$

In quantum mechanics the superposition state represents a conflicting situation. According to the coherence principle the state

$$\alpha|0\rangle + \beta|1\rangle$$

is a new state where the single states $|0\rangle$ and $|1\rangle$ are inseparable from each other: we cannot reduce the superposition states to the elementary states $|0\rangle, |1\rangle$. This is in agreement with the evidence theory where the evidence for every subset is independent from the evidence of the elementary states

$$P_1 = m(\{1\}) = |\psi_1(x)|^2 = \int_{\Omega} \psi_1(x)\bar{\psi}_1(x)\,\mathrm{d}x = \langle 1|1\rangle,$$

$$P_0 = m(\{0\}) = |\psi_0(x)|^2 = \int_{\Omega} \psi_0(x)\bar{\psi}_0(x)\,\mathrm{d}x = \langle 0|0\rangle.$$

Because the superposition state is $\psi = \alpha|0\rangle + \beta|1\rangle$ we have

$$\begin{aligned}\langle\psi|\psi\rangle = m(\{1,0\}) = |\psi(x)|^2 = |\alpha\psi_1(x) + \beta\psi_0(x)|^2 \\ = \alpha^2[\psi_1(x)\cdot\psi_1(x)] + \beta^2[\psi_0(x)\cdot\psi_0(x)] + 2\alpha\beta[\psi_0(x)\cdot\psi_1(x)],\end{aligned}$$

where "·" denotes the scalar product $\int_\Omega \psi(x)\bar{\psi}(x)\,\mathrm{d}x$

$$\psi_0(x)\cdot\psi_1(x) = \int_\Omega \psi_0(x)\bar{\psi}_1(x)\,\mathrm{d}x\,\mathrm{d}y.$$

So, we have

$$\begin{aligned}&m(\{0\}) = \alpha^2[\psi_1(x)\cdot\psi_1(x)], \quad m(\{1\}) = (\beta^2[\psi_0(x)\cdot\psi_0(x)],\\ &m(\{0,1\}) = m(\{0\}) + m(\{1\}) + 2\alpha\beta[\psi_0(x)\cdot\psi_1(x)],\end{aligned}$$

where $2\alpha\beta[\psi_0(x,y)\cdot\psi_1(x,y)]$ is the interference term that can be positive or negative.

From the theory of evidence [18], we know that the evidence is connected with any subset without any a priori restriction or connection between the elements of the universal set and the subsets. We are completely free to give any type of evidence to the subsets. If A and B are disjoint sets, $A \cap B = \emptyset$, when we know the evidence of A and B (basic probability assignment) we cannot use any type of rules to calculate the evidence of $m(A \cup B)\cdot m(A)$, $m(B)$ and $m(A \cap B)$ can assume any value.

So we can distinguish three possible cases:

1. $m(A \cup B) > m(A) + m(B)$ superadditivity, the evidence of the union is greater of the sum of the evidence for the single sets;
2. $m(A \cup B) = m(A) + m(B)$ additivity, the evidence of the union is equal of the sum of the evidence for the single sets;
3. $m(A \cup B) < m(A) + m(B)$ subadditivity, the evidence of the union is less of the sum of the evidence for the single sets.

When we have $m(A \cup B) = m(A) + m(B)$ (additivity case), we have the traditional probability rule for $A \cap B = \emptyset$.

With the quantum mechanical representation we can suggest a new vectorial interpretation of the basic probability assignment. In fact when we associate with the singleton sets $\{1\}$ and $\{0\}$ the infinite dimensional vectors $\psi_1(x)$, $\psi_0(x)$ (the wave function for the qubits $|1\rangle$ and $|0\rangle$, respectively)

$$\{1\} \to \psi_1(x), \ \{0\} \to \psi_0(x)$$

in the formula $m(\{0,1\}) = m(\{0\}) + m(\{1\}) + S_{01}$.

S_{01} is given by the expression

$$S_{01} = 2\alpha\beta[\psi_0(x)\cdot\psi_1(x)].$$

We can denote the term S_{01} *synergetic term* between the state $|1\rangle$ or "true" and the state $|0\rangle$ or "false".

In this way we obtain a simplification of the evidence theory and it is exactly this observation that justifies the intuitive ideas of sensor fusion [19]. Given a universal set U we can associate with each subset with only one element (singleton) a wave function or a vector. The evidence of the singletons is proportional to the scalar product of each vector with itself. The evidence of the other subsets is proportional to the mixed scalar products.

10.1. The fundamental tensor $g_{i,j}$ and the difference between probability and evidence

Previously we used the fundamental tensor $g_{i,j}$ to help us discover models for morphogenetic filters in a non-orthogonal functional space. In this section we will show that the same fundamental tensor $g_{i,j}$ can be used to create a vector model of evidence theory. We discover that the $g_{i,j}$ gives us the vehicle to compute the synergy between elements in the set. With this tensor we reduce the information necessary to know the basic probability assignment from 2^N sets to N vectors. In this way we give a geometrical meaning of evidence theory, in a similar way as indeterminacy in quantum mechanics is formalized by the geometry of Hilbert spaces. Quantum mechanics inspired us to this new model of evidence theory. The model presented in this section is not complete. In the next section we suggest an extension of the scalar product to improve the model.

Given the universal set $U = \{a, b, c\}$, wave functions or vectors $\psi_a(x),\ \psi_b(x),\ \psi_c(x)$ associate to the singletons

$$\{a\} \to \psi_a(x),\ \ \{b\} \to \psi_b(x),\ \ \{c\} \to \psi_c(x).$$

The evidence of the universal set is

$$\begin{aligned} m(U) = m(\{a,b,c\}) &= k(\psi_a(x) + \psi_b(x) + \psi_c(x)) \cdot (\psi_a(x) + \psi_b(x) + \psi_c(x)) \\ &= k \sum_{i,j} \psi_i \cdot \psi_j = \sum_{i,j} g_{i,j}, \end{aligned}$$

where

$$g_{i,j} = G = \begin{bmatrix} \psi_a \cdot \psi_a & \psi_a \cdot \psi_b & \psi_a \cdot \psi_c \\ \psi_b \cdot \psi_a & \psi_b \cdot \psi_b & \psi_b \cdot \psi_c \\ \psi_c \cdot \psi_a & \psi_c \cdot \psi_b & \psi_c \cdot \psi_c \end{bmatrix} = \begin{bmatrix} m(\{a\}) & S_{a,b} & S_{a,c} \\ S_{b,a} & m(\{b\}) & S_{b,c} \\ S_{c,a} & S_{c,b} & m(\{c\}) \end{bmatrix}$$

so

$$m(\{a,b,c\}) = m(\{a\}) + m(\{b\}) + m(\{c\}) + \sum(\text{synergetic terms}),$$

when the vectors $\psi_a(x), \psi_b(x), \psi_c(x)$ are orthogonal we have

$$m(\{a,b,c\}) = m(\{a\}) + m(\{b\}) + m(\{c\})$$

and the evidence theory reduces to classical probability theory (additivity theory).

Example 3. Let the universal set be $U = \{(0,0),(1,0),(0,1),(1,1)\}$.

States $|00\rangle,\ |10\rangle,\ |01\rangle,\ |11\rangle$ superposition

$$\psi = \alpha_1|00\rangle + \alpha_2|10\rangle + \alpha_3|01\rangle + \alpha_4|11\rangle.$$

Evidence value for the universal set U

$$m(U) = [\alpha_1|00\rangle + \alpha_2|10\rangle + \alpha_3|01\rangle + \alpha_4|11\rangle] \cdot [\alpha_1|00\rangle + \alpha_2|10\rangle + \alpha_3|01\rangle + \alpha_4|11\rangle]$$

for

$$m(\{00\} = \alpha_1^2\langle 00|00\rangle = \alpha_1^2\psi_{00} \cdot \psi_{00}, \quad m(\{01\} = \alpha_2^2\langle 01|01\rangle = \alpha_2^2\psi_{01} \cdot \psi_{01},$$
$$m(\{10\} = \alpha_3^2\langle 10|10\rangle = \alpha_3^2\psi_{10} \cdot \psi_{10}, \quad m(\{11\} = \alpha_4^2\langle 11|11\rangle = \alpha_4^2\psi_{11} \cdot \psi_{11},$$

we have $m(U) = m(\{a\}) + m(\{b\}) + m(\{c\}) + \sum$(synergetic terms), where the synergetic terms are given by the matrix

$$g_{i,j} = \begin{bmatrix} \psi_{00}\cdot\psi_{00} & \psi_{00}\cdot\psi_{01} & \psi_{00}\cdot\psi_{10} & \psi_{00}\cdot\psi_{11} \\ \psi_{01}\cdot\psi_{00} & \psi_{01}\cdot\psi_{01} & \psi_{01}\cdot\psi_{10} & \psi_{01}\cdot\psi_{11} \\ \psi_{10}\cdot\psi_{00} & \psi_{10}\cdot\psi_{01} & \psi_{10}\cdot\psi_{10} & \psi_{10}\cdot\psi_{11} \\ \psi_{11}\cdot\psi_{00} & \psi_{11}\cdot\psi_{01} & \psi_{11}\cdot\psi_{10} & \psi_{11}\cdot\psi_{11} \end{bmatrix}.$$

When $\sum$(synergetic terms) > 0 we have *positive synergy* and

$$m(U) > m(\{a\}) + m(\{b\}) + m(\{c\}).$$

The evidence of U is greater of the evidence of the parts. The parts are semantically more uncertain when $\sum$(synergetic terms) $= 0$ we have *zero synergy* and

$$m(U) = m(\{a\}) + m(\{b\}) + m(\{c\}).$$

The evidence of U is equal of the evidence of the parts. The evidence in this case is equal to the probability. The fundamental tensor is $g_{i,j} = \delta_{i,j}$ *the functions are orthogonal.*

When $\sum$(synergetic terms) < 0 we have *negative synergy* and

$$m(U) = m(\{a\}) + m(\{b\}) + m(\{c\}).$$

The evidence of U is less of the evidence of the parts. The parts are semantically less uncertainly that the universe.

11. Extension of the scalar product

From our discussions in the previous sections we have seen that the scalar product plays a pivotal role in the biology of pattern recognition. If one studies the measurements in Fig. 4 carefully, one observes that although the desired feature clearly provokes by far the largest neuron response, especially in the case that the stimulus is coherent, still the other features show non-zero response. This phenomenon can of course be partially attributed to measurement errors, but it is interesting to note that when one extends the classical scalar product to a more general construct, the "leakage" of information is a general consequence of the scalar product extension. We try to explain our ideas in the following.

The scalar product appears as too simple.

Geometrically, the extension of the scalar product can be thought of as one of the two vectors influencing the other before taking the classical scalar product (i.e. a local space deformation). The generalization of the classical scalar product to a more general concept has the consequence that we assume some coherence between the two vectors and this can in turn be interpreted as a "softening" of the vector independence: we cannot separate one vector from the other. This can also be interpreted as uncertainty or fuzziness in the definition of the vectors. The generalization of the scalar product used in this paper is only one of the possible models. By introducing this new scalar product, we move away from the conventional concepts used in Hilbert space, particularly in quantum mechanics. Since quantum mechanics is not the only vehicle to model interpretation of knowledge, we suggest a more flexible instrument, closely resembling the cognitive approach to knowledge.

We may compare this with the existence of a family of t-norms and t-conorms to model operations between fuzzy variables. The existence of such families is justified by the need for a general mechanism to model human intuitive logic operations used in cognitive inference, since classical propositional logic is too rigid for this. In a similar way the generalized scalar product induces flexibility in the computation of the morphogenetic neuron described before. This is more in line with recent findings on the functioning of the brain [20]. Additional evidence for the need of a more general scalar product concept follows from evidence theory: for every member of the power set we can assign arbitrary values to the basic probability assignment. With the scalar product we reduce the information to single elements and with some scalar product rules we come back to the power set. This mechanism cannot rebuild completely all possible situations in evidence theory, such as total ignorance. Only the extension of scalar product can enlarge the possibilities offered by vectorial and scalar product approach. This is the fundamental motivation for extension of the scalar product. Quantum mechanics on itself cannot account for the human cognitive approach to uncertainty by evidence theory. Extension of the scalar product helps to introduce a substantial improvement to the human approach to uncertainty studied in evidence theory. We introduce in the following one of the possible models for the scalar product extension. We remember that any extension of the scalar product must satisfy at the orthogonality property and the boundary condition.

Let a two-dimensional Euclidean space (x, y) be given. We can then define states in this space as follows:

$$\psi_1(x,y) = |1\rangle, \psi_2(x,y) = |2\rangle, \ldots, \psi_n(x,y) = |n\rangle.$$

If we take Eq. (2) with $w_{11} = 1,\ w_{12} = 1,\ w_{21} = 1$ and $w_{22} = -1$ we have the two morphogenetic neurons:

$$u_1 = |+\rangle = |i\rangle + |k\rangle \quad \text{and} \quad u_2 = |-\rangle = |i\rangle - |k\rangle.$$

The probabilities of the mixed states are

$$\langle +|+\rangle = \int [\psi_i(x,y) + \psi_k(x,y)]^2 \,\mathrm{d}x\,\mathrm{d}y \quad \text{and}$$

$$\langle -|-\rangle = \int [\psi_i(x,y) - \psi_k(x,y)]^2 \,\mathrm{d}x\,\mathrm{d}y.$$

The output function of the morphogenetic neuron is

$$u_{j,k} = \sum_{i=1}^{n} w_{j,i} \langle i|k\rangle_G, \tag{9}$$

where

$$\langle i|k\rangle_G = \exp(-\alpha\langle -|-\rangle) - \exp(-\alpha\langle +|+\rangle). \tag{10}$$

The WRITE operation consists in the computation from (9) of the weight parameters $w_{j,i}$. To understand the meaning of Eq. (10) we will expand it in Taylor form as follows:

$$\begin{aligned}\langle i|k\rangle_G &= 4\alpha \int_\Gamma \psi(x,y)_i \psi(x,y)_k \,\mathrm{d}x\,\mathrm{d}y \left[1 - \alpha \int_\Gamma (\psi(x,y)_i)^2 \,\mathrm{d}x\,\mathrm{d}y \right. \\ &\qquad \left. + \int_\Gamma (\psi(x,y)_k)^2 \,\mathrm{d}x\,\mathrm{d}y + \cdots \right] \\ &= 4\alpha\langle i|k\rangle[1 - \alpha(\langle i|i\rangle + \langle k|k\rangle + \cdots],\end{aligned}$$

where $\int_\Gamma \psi(x,y)_i \psi(x,y)_k \,\mathrm{d}x\,\mathrm{d}y = \langle i|k\rangle$ is the scalar product of the states $\langle i|$ and $|k\rangle$. So $\langle i|k\rangle_G$ is an *extension* of the scalar product. When the parameter $\alpha \to 0$ the extension of the scalar product reduces to the ordinary scalar product $\langle i|k\rangle$ see e.g. [8]. If α is increased from zero we create a continuous set of operations parameterized with α. In this way we have a flexible definition of the scalar product that can be used for different filter operations, thus extending the classical projection (filter) operator.

12. Conclusions

In this publication we have applied the powerful concept of the morphogenetic neural network architecture proposed previously by us and have combined this concept with the theory of quantum computation. We have described a procedure using a morphogenetic filter that is capable of extracting invariants and abstract rules from arbitrary numerical data sets. The procedure proves to be very efficient and can be used as an alternative to extract minimum expressions depending on the type of data, completely eliminating cumbersome and error-prone calculus and analysis. The invariants generally are found in implicit form. They allow for multi-valued solutions and are therefore more powerful than explicit formulae.

An example of this technique is given: solving (inverse) coordinate transformations in robotics without symbolic manipulation.

Finally we discovered how to give a novel model of evidence theory based on geometry of the Hilbert space. With this model we reduce the power set of U by a set U^* of vectors that is isomorphic to U. With the set U^* with the scalar product or extension of the scalar product we can generate the value of the basic probability assignment for all sets in the power set. We remark that quantum superposition in Hilbert space has inspired these ideas to use the MN to evidence theory.

In future we intend to look for possibilities to map the algorithm on a quantum computer and to demonstrate the capability of the MN to realize massive parallel computation until now only expected for quantum computing.

Acknowledgements

One of the authors (AJvdW) gratefully acknowledges the financial support by Strukton Systems B.V.

References

[1] W. McCullough, W.H. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (1943) 115–133.
[2] J.M.J. Murre, Learning and Categorization in Modular Neural Networks, Erlbaum, Hillsdale, NJ, 1992.
[3] C. Koch, Computation and the single neuron, Nature 385 (16) (1997) 207–210.
[4] J.G. Benjafield, Cognition, Prentice-Hall, Englewood Cliffs, NJ, 1992.
[5] G. Resconi, in: O. Kaynak, L. Zadeh, B. Turksen, I.J. Rudas (Eds.), The Morphogenetic Neuron in Computational Intelligence: Soft Computing and Fuzzy – Neuro Integration with Application, Springer NATO ASI Series F Computer and System Science, vol. 162, 1998, pp. 304–331.
[6] G. Resconi, E. Pessa, R. Poluzzi, in: The Morphogenetic Neuron, Proceedings of the Fourteenth European Meeting on Cybernetics and Systems Research, April 14–17, 1998, pp. 628–633.
[7] G. Resconi, A.J. van der Wal, A data model for the Morphogenetic Neuron, Int. J. General System 29 (1) (2000) 141–174.

[8] G. Resconi, A.J. van der Wal, Single neuron as a quantum computer-morphogenetic neuron, in: Intelligent techniques and Soft Computing in Nuclear Science and Engineering, Proceeding of the 4th International FLINS Conference Bruges, Belgium, August 28–30, World Scientific, Singapore, 2000.

[9] A. Barenco, Quantum computation, Thesis Lincoln College University of Oxford Trinity Term, 1996.

[10] G. Brassard, I. Chuang, S. Loyd, C. Monroe, Quantum computing, Proc Nat. Acad. Sci. USA 95 (September) (1998) 11032–11033.

[11] P.W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A 52 (1995) 2493–2496.

[12] S. Nobuo, The extent to which Biosonar information is represented in the Bat Auditory Cortex, in: G.M. Edelman (Ed.), Dynamic Aspects of Neocortical Function, John Wiley, New York, 1984.

[13] E. Salinas, L.F. Abbott, A model of multiplicative neural responses in parietal cortex, Proc. Natl. Acad. Sci. USA 93 (1996) 11956–11961.

[14] E. Salinas, L.F. Abbott, in: Invariant Visual responses From Attentional gain Fields, The American Physiological Society, 1997, pp. 3267–3272.

[15] A.I. Borisenko, I.E. Tarapov, Vector and Tensor Analysis with Applications, Prentice-Hall, Englewood Cliffs, NJ, 1968.

[16] A.J. van der Wal, The role of fuzzy set theory in the conceptual foundations of quantum mechanics: an early application of fuzzy measures, in: Foundations and applications of possibility theory, Advances in Fuzzy Systems, FAPT 95, vol. 8, World Scientific, Singapore, 1995, pp. 234–245.

[17] G. Resconi, G. George, Klir, E. Pessa, Conceptual foundation of quantum mechanics: the role of evidence theory, quantum sets, and modal logic, Int. J. Modern Phys. C 10 (1) (1999) 29–62.

[18] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.

[19] A.J. van der Wal, Sensor synergetics: the design philosophy of the sensor fusion demonstrator and testbed at TNO-FEL, in: Multi-Sensor Systems and Data Fusion for Telecommunications, Remote Sensing and Radar, AGARD-CP-595, Lisbon, vol. 31, 1998, pp. 1–26 (ISBN 92-836-0051-17).

[20] J.D. Walls, K.C. Anderson, E.K. Miller, Single neurons in prefrontal cortex encode abstract rules, Nature 41/21 (June) (2001).

[21] A. Pouget, L.H. Snyder, Computational approaches to sensorimotor transformation, Nature Neurosci., 3 (Suppl.) (2000) 1192–1198.

[22] M. Riesenhuber, T. Poggio, Models of object recognition, Nature Neurosci. 3 (Suppl.) (2000) 1199–1204.