Studies in Computational Intelligence

Krzysztof Pancerz
Elena Zaitseva *Editors*

Computational Intelligence, Medicine and Biology

Selected Links

Springer

Studies in Computational Intelligence

Volume 600

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Studies in Computational Intelligence" (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

More information about this series at http://www.springer.com/series/7092

Krzysztof Pancerz · Elena Zaitseva
Editors

Computational Intelligence, Medicine and Biology

Selected Links

Springer

Editors
Krzysztof Pancerz
Department of Computer Science and Knowledge Engineering
University of Management and Administration
Zamosc
Poland

and

Institute of Biomedical Informatics
University of Information Technology and Management
Rzeszow
Poland

Elena Zaitseva
Department of Informatics
University of Zilina
Zilina
Slovakia

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-319-16843-2 ISBN 978-3-319-16844-9 (eBook)
DOI 10.1007/978-3-319-16844-9

Library of Congress Control Number: 2015937047

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

Computational intelligence plays an important role in different areas of human activity. Reasoning, prediction, modeling, optimization, decision making, etc. need modern, soft and intelligent algorithms, methods and methodologies to solve, in the efficient ways, different real life problems encountered by humans, among others, in the following fields: economy, finance, security, sociology, psychology, biology, medicine and health care. There are a lot of research groups over the world attempting to develop modern tools utilizing computational intelligence in these fields. The idea of the book titled *Computational Intelligence, Medicine and Biology: Selected Links* is to collect several examples of such attempts. The attention is focused on two fields, namely, medicine and health care as well as biology. The preparation of the book was mainly possible thanks to cooperation of editors and some of the authors under different research projects, e.g., the grant No. SK-PL-0023-12 under *Research and Development Cooperation Slovakia-Poland* and *Physarum Chip Project: Growing Computers from Slime Mould* (PhyChip) funded by the Seventh Framework Programme (FP7) as well as on the basis of selected contributions to the International Workshop on Artificial Intelligence in Medical Applications (AIMA'2013) organized in conjunction with the Federated Conference on Computer Science and Information Systems (FedCSIS'2013) held in September 2013 in Kraków, Poland. We appreciate all contributions very much.

The contents of the book were divided into two parts:

- Part I. Computational Intelligence and Medicine.
- Part II. Computational Intelligence and Biology.

Part I, consisting of four chapters, is devoted to selected links of computational intelligence, medicine, health care and biomechanics. Reliability is one of the most important characteristics of any health care system. E. Zaitseva et al. describe a new method for the estimation of healthcare system reliability. This method allows the users to analyze principal components of healthcare systems as devices, equipment and human factors. Z. Omiotek et al. discuss results of experiments with classification of ultrasound thyroid images using selected machine learning methods. T. Nawarycz et al. show application of fuzzy logic to measure weight status and central

fatness in adults and adolescents. The presented solution is useful in the assessments of the cardio-metabolic risks. The last chapter of Part I, given by J. Majernik, concerns a new approach to deriving kinematics directly from video records where no special apparel is needed. In the proposed approach, the digital image processing techniques are used to develop algorithms for automatic detection of human motion trajectories.

Part II, also consisting of four chapters, is devoted to selected links of computational intelligence and biology. The common denominator of three chapters is *Physarum polycephalum*, one-cell organisms able to build complex networks for solving different computational tasks. The behavior of collectives of trematode larvae (miracidia and cercariae) can be considered adaptable to environmental changes. For describing that behavior, A. Schumann and L. Akimova appeal to a language which is a kind of process calculus called *Physarum* spatial logic that is based on a formalization of the behavior of *Physarum polycephalum*. The chapter given by J.D. Jones concerns the approximation of Lateral Inhibition phenomena in unorganised, non-neural systems, using a model of slime mould *Physarum polycephalum*. This chapter addresses the question of how simple organisms possessing no nervous system might approximate collective signal enhancement analogous to Lateral Inhibition using bulk transport of material. Evacuation is an imminent movement of people away from sources of danger. Finding a shortest path towards exit is a key for the prompt successful evacuation. V.S. Kalogeiton et al. show that slime mould *Physarum polycephalum* is proven to be an efficient path solver. The last chapter, given by E. Gale et al., focuses on a novel device, the memristor, that has possible uses both in the creation of hardware neural nets for artificial intelligence and as the connection between a hardware neural net and a living neuronal cell network for uses in the treatment and monitoring of neurological disease.

We hope that all chapters will find a wide audience of readers who will consider them interesting.

January 2015

Krzysztof Pancerz
Elena Zaitseva

Contents

Part II: Computational Intelligence and Biology

Part I
Computational Intelligence and Medicine

Estimation of a Healthcare System Based on the Importance Analysis

Elena Zaitseva, Miroslav Kvassay, Vitaly Levashenko,
Jozef Kostolny, and Krzysztof Pancerz

Abstract. Most of modern systems are complex and include heterogeneous subsystems and components, such as software, devices, equipment, that are influenced by human factor. Healthcare systems are a typical example of such systems that have to guarantee the perfect care of a patient and have high reliability. Therefore, the reliability analysis of a healthcare system is an important problem. There are different methods in reliability engineering for analysis and quantification of every healthcare component, but a tendency in reliability analysis needs new methods that evaluate the system as a whole. In accordance with this tendency, one aspect of reliability engineering, known as the importance analysis, is considered. The importance analysis is one of the approaches in reliability engineering. In this paper, new algorithms for the calculation of importance measures are developed. These measures permit to estimate the influence of any component (sub-system) on a healthcare system functioning or failure.

Elena Zaitseva · Miroslav Kvassay · Vitaly Levashenko · Jozef Kostolny
Department of Informatics, University of Zilina, Zilina, Slovakia
e-mail: {jozef.kostolny,miroslav.kvassay,
vitally.levashenko,elena.zaitseva}@fri.uniza.sk

Krzysztof Pancerz
Department of Computer Science and Knowledge Engineering,
University of Management and Administration, Zamość, Poland
e-mail: kpancerz@wszia.edu.pl

Institute of Biomedical Informatics,
University of Information Technology and Management, Rzeszów, Poland

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*,
Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_1

1 Introduction

There is one principal characteristic for all healthcare systems. It is reliability that is defined as the probability that a system will perform its intended function during a period of running time without any failure. The system reliability is a complex characteristic that depends on the functioning of separate parts (components) of the system and depends on the combination of a number of interrelated processes of every component as degradation, failure and repair, which result from the interaction of different parts including not only the hardware but also the software, the human and the organizational system parts.

Application of reliability engineering methods for healthcare system investigation was proposed more than 30 years ago in paper [1]. E. F. Taylor declared principal items of reliability engineering in a healthcare system as the reliability analysis of medical equipment and devices. Reliability quantification of equipment and devices has been a principal tendency in medicine recently. Information technology development causes application of a new type of healthcare systems that consist of two interdependent components such as hardware and software [2, 3]. Therefore, new methods for the shared analysis of the hardware and software part of a healthcare system have been developed. One of these methods was presented in paper [4], but a human error problem in a healthcare system was considered as an independent problem of the reliability analysis [5, 6].

Based on bibliography of the reliability analysis of the healthcare domain, we can show two principal approaches to reliability engineering in medicine. The first one is reliability estimation of medical equipment and devices that includes reliability quantification of hardware and software of the healthcare system [3, 4, 7]. The second approach agrees with examination of human errors [5, 6, 8]. However, independent evaluation of these principal parts of the healthcare system is not allowed to have detailed and actual reliability analysis. In paper [9], new tendencies in reliability engineering are considered. According to [9], the reliability analysis has to be based on joint evaluation of all principal parts (components).

The typical structure of the healthcare system consists of some principal components from the point of view of the reliability analysis [9, 10, 11]. In papers [3, 11], two of them have been defined: equipment/device and human factor. We need to note that the human factor was considered as errors of operators of medical equipment or devices in [3]. A detailed structure of the human factor and human errors for the healthcare system is presented in [11]. The healthcare system structure includes three components: technical, human and organizational [9]. The technical component includes two types of medical devices/equipment that are based on special and standard technologies according to [9, 10]. For example, the first type is a medical decision support system, the system for integration of electronic medical records or the picture archiving communication system. The second type is a special medical device with equipment that can be used for a special operation only (for example, magnetic resonance imaging scanners). The human component of the healthcare system models medical errors. The organizational component of the system joins management aspects and maintenance of the healthcare system.

The important problem in the reliability analysis of the healthcare system is the development of methodology that permits to investigate every system component and the system on the basis of unified approaches. This conception was presented in papers [9, 12]. We propose and develop one of the possible approaches for investigation of healthcare system reliability. It is the importance analysis that allows us to estimate the influence of every system component state change on the system performance (reliability/availability). The initial system in such analysis is declared as a Multi-State System (MSS). MSS is a mathematical model in the reliability analysis that represents system performance behaviour as the reliability with some level: from complete failure to perfect functioning. This interpretation of a system allows us to investigate different performance levels of system functioning, that do not include the level of functioning and fault only.

Direct Partial Logic Derivatives [13] are the mathematical background of the MSS importance analysis. The advantage of this approach is the possibility of using it for estimation of every system and its components. For example, in paper [9], the application of this approach for the estimation of a healthcare system of four components was considered. In this paper, new algorithms for calculation of the importance measures, such as Birnbaum importance and Fussell-Vesely importance, are developed.

The paper is organized as follows. In the following Section, the basic conception of the healthcare system reliability analysis is introduced. Then, Section 3 presents a mathematical model of a healthcare system based on MSS and new algorithms for the calculation of the importance measures. A numerical example for the human module as a part of the healthcare system is provided in this section too.

2 Reliability Analysis of a Healthcare System

The reliability analysis of a system includes three principal steps [10]:

- the quantification of the system model;
- the representation and modelling of the system;
- the representation, propagation and quantification of the uncertainty in the system behavior.

Let us consider characteristics of these steps for the healthcare system in more details.

2.1 Quantification of the System

Definition of the number of performance levels for the system in the reliability analysis is a principal step that influences the development of a mathematical model too. As a rule, two approaches are used for the quantification in reliability engineering.

The first of them defines only two states for the system reliability: the functioning and failure. The mathematical model for the representation of this quantification is called a Binary-State System (BSS). This approach is well known and widely used in reliability engineering. The system failure can be investigated in detail on the basis of this quantification. However, the analysis of other performance levels before the system failure has some difficulties for BSS. In this case, the quantification of the system reliability to some performance levels is used. The mathematical model with more than two performance levels is called a Multi-State System (MSS).

The MSS reliability analysis is a more flexible approach to evaluating system reliability, as it can be used when both the system and its components may experience more than two states, including, for example, completely failed, partially failed, partially functioning and perfectly functioning [14, 15]. However, a mathematical approach to analysis of such a system is complex.

2.2 Modelling of the System

The next step, after the definition of quantification, is the mathematical model development. There are some types of the BSS representations as the mathematical model. These representations (mathematical models) correlate with the mathematical methods for the calculation of the system reliability indices and measures. For example, the description of the system by the Universal Generating Function is used in the system reliability optimization [16]. Another representation is the structure function. This function allows us the mathematical description of a system with any complexity [13, 14].

The *structure function* defines the system state (system reliability/performance level) depending on the system component states. According to the definition of the structure function, the system reliability in the stationary state is defined as [13]:

$$\phi(x_1,...,x_n) = \phi(\mathbf{x}) : \{0,...,m_1-1\} \times ... \times \{0,...,m_n-1\} \rightarrow \{0,...,M-1\}. \quad (1)$$

In (1), x_i is the state of the i-th system component that can be defined from 0 (the component failure) to $m_i - 1$ (the perfect component performance level), i.e., $x_i \in \{0,...,m_i-1\}$, and the system reliability has M levels from 0 (as the failure) to $M-1$ (as the perfect functioning). It is worth noting that the system components have different numbers of states: $m_i \neq m_k$, if $i \neq k$ $(i,k \in \{1,...,n\})$. The number of the system components is declared as n.

The structure function definition (1) is the definition for MSS, where the system and its components have some (more than two) performance levels. The structure function for BSS is defined as a Boolean function [17]:

$$\phi(x_1,...,x_n) = \phi(\mathbf{x}) : \{0,1\}^n \rightarrow \{0,1\}. \quad (2)$$

We will consider a coherent system in this paper. Such a system has two principal assumptions for the structure function [14]: (a) the structure function (1) or (2) is

monotone, and (b) the system component state decrease does not improve the system reliability. Every system component is characterized by probability of its states:

$$p_{i,s_i} = \Pr\{x_i = s_i\}, s_i \in \{0,...,m_i - 1\}. \tag{3}$$

Let us note, the structure function (1) or (2) presents the investigated object (system) in the stationary state. The system behavior and correlation of changes of component states and system reliability can be defined by mathematical tools of Logical Differential Calculus, in particular, the Direct Partial Logic Derivative. The Direct Partial Logic Derivative with respect to variable x_i for the structure function (1) permits to analyse the system reliability change from j to $\overline{j}$ when the i-th component state changes from s to $\overline{s}$ [13, 17]:

$$\frac{\partial \phi(j \to \overline{j})}{\partial \phi x_i(s \to \overline{s})} = \begin{cases} 1 \text{ if } \phi(s_i, \mathbf{x}) = j \text{ and } \phi(\overline{s}_i, \mathbf{x}) = \overline{j}, \\ 0 \text{ otherwise,} \end{cases} \tag{4}$$

where:

- $\phi(s_i, \mathbf{x}) = \phi(x_1,...,x_{i-1}, s, x_{i+1},...,x_n)$, $\phi(\overline{s}_i, \mathbf{x}) = \phi(x_1,...,x_{i-1}, \overline{s}, x_{i+1},...,x_n)$,
- $s, \overline{s} \in \{0,...,m_i - 1\}$, $s \neq \overline{s}$,
- $j, \overline{j} \in \{0,...,M-1\}$, $j \neq \overline{j}$.

Let us consider the system failure in the Direct Partial Logic Derivative terminology. The system failure is represented as a change of the structure function value $\phi(\mathbf{x})$ from state 1 to 0. This change can be caused by the i-th variable change from 1 to 0 if we consider a coherent system. Therefore, the Direct Partial Logic Derivative for the system failure analysis is defined by the equation:

$$\frac{\partial \phi(1 \to 0)}{\partial \phi x_i(s \to s-1)} = \begin{cases} 1 \text{ if } \phi(s_i, \mathbf{x}) = 1 \text{ and } \phi((s-1)_i, \mathbf{x}) = 0, \\ 0 \text{ otherwise.} \end{cases} \tag{5}$$

This derivative for the system with two performance levels (BSS) is defined as:

$$\frac{\partial \phi(1 \to 0)}{\partial \phi x_i(1 \to 0)} = \begin{cases} 1 \text{ if } \phi(1_i, \mathbf{x}) = 1 \text{ and } \phi(0_i, \mathbf{x}) = 0, \\ 0 \text{ otherwise.} \end{cases} \tag{6}$$

The Direct Partial Logic Derivative (5) allows investigating boundary states of this system for which reduction of one component x_i causes the system failure. The Direct Partial Logic Derivative (6) describes the situation for which the system fails depending on one system component failure. Therefore, these derivatives allow us to calculate the system boundary states for the system reliability that agrees with vector state $\mathbf{x} = (x_1, x_2, ..., x_n)$ and the change of one of the component states causes the change of the system performance level (system reliability).

There is one more type of the Direct Partial Logic Derivative that allows us to investigate the influence of all changes of the component to the system reliability. It is a union of all derivatives $\partial \phi(j \to \overline{j}) / \partial x_i(s \to s+1)$. This union of derivatives identifies all situations, in which the improvement of the fixed component state results in the transition of the system from the state less than j to the state greater than or equal to j:

$$\frac{\partial \phi(\uparrow j \uparrow)}{\partial x_i(s \to s+1)} = \bigcup_{w=0}^{j-1} \left(\bigcup_{v=j}^{M-1} \frac{\partial \phi(w \to v)}{\partial x_i(s \to s+1)} \right), \tag{7}$$

for $s = 0,...,m_i - 2$ and $j = 1,...,M-1$.

The Direct Partial Logic Derivative union (7) can be defined as:

$$\frac{\partial \phi(\uparrow j \uparrow)}{\partial x_i(s \to s+1)} = \begin{cases} 1 \text{ if } \phi(s_i, \mathbf{x}) < j \text{ and } \phi((s+1)_i, \mathbf{x}) \geq j, \\ 0 \text{ otherwise} \end{cases} \tag{8}$$

for $s = 0,...,m-2$ and $j = 1,...,m-1$.

The dimension of the Direct Partial Logic Derivative with respect to the variable x_i is defined as m^{n-1} [20]. Therefore, the Direct Partial Logic Derivative union (7) and (8) for the i-th variable has dimension m^{n-1} too, but the structure function investigation supposes the analysis of all the possible system states. Therefore, the Direct Partial Logic Derivative union (7) or (8) has to be transformed for dimension m^n as follows:

$$\frac{\partial \phi(\uparrow j \uparrow)}{\partial x_i(s \to s+1)} = \begin{cases} 1 \text{ if } x_i = s \text{ and } \phi(s_i, \mathbf{x}) < j \text{ and } \phi((s+1)_i, \mathbf{x}) \geq j, \\ 0 \text{ if } x_i = s \text{ and } \phi(s_i, \mathbf{x}) \geq j \text{ or } \phi((s+1)_i, \mathbf{x}) < j, \\ * \text{ if } x_i \neq s. \end{cases} \tag{9}$$

The merge (union) of unions (9) of Direct Partial Logic Derivatives has to be computed to identify all situations, in which any improvement of a given component causes the transition of the system from the state less than j to the state greater than or equal to j:

$$\frac{\partial \phi(\uparrow j \uparrow)}{\partial x_i} = \bigcup_{s=0}^{m-2} \frac{\partial \phi(\uparrow j \uparrow)}{\partial x_i(s \to s+1)} \text{ for } j = 1,2,...m-1. \tag{10}$$

The intersection of two merges (10) of two different variables (components) is defined in Table 1. This intersection identifies state vectors, in which the change of states of both components (if the component state can be changed) results in the change of the system state from the value less than j to the value greater than or equal to j. In Table 1, symbol "1" means that the state of at least one component from components i_1 and i_2 can be changed and all those changes result in the required change of the system state (from the state less than j to the state greater than or equal to j). Symbol "0" identifies those state vectors, in which at least one component change does not cause the required change of the system state. Finally, symbol "$*$" correlates with those situations, when neither component i_1 nor component i_2 can be changed from state s to state $s+1$, because component i_1 is in state $m_{i_1} - 1$ and component i_2 is in state $m_{i_2} - 1$.

The intersection of two merges of extended unions of Direct Partial Logic Derivatives identifies state vectors, in which the improvement of both components (if components can be repaired, i.e., if component i_1 is not in state $m_{i_1} - 1$ and component

Table 1 The intersection of two merges of unions of Direct Partial Logic Derivatives

		$\frac{\partial\phi(\uparrow j \uparrow)}{\partial x_{i_2}}$		
		*	**0**	**1**
$\frac{\partial\phi(\uparrow j \uparrow)}{\partial x_{i_1}}$	*	*	0	1
	0	0	0	0
	1	1	0	1

i_2 is not in state $m_{i_2} - 1$) results in the improvement of the system from the state less than j to the state greater than or equal to j.

The merge (10) of extended unions of Direct Partial Logic Derivatives has to be computed for all components of the system and then their "intersection" allows us to identify *Minimal Cut Sets* (MCSs) for system state j. MCS is one more type of the boundary states of the system for which the component state improvement causes the system reliability change. MCS is one of the basic conceptions in reliability engineering that is used to calculate a lot of reliability measures and indices.

Let us consider the conception of the minimal cut set (MCS). At first, MCS methods have been proposed for the BSS reliability analysis. According to [18, 19] a cut set is defined as a set of components of a system whose simultaneous failure leads into the failure of the system (if the system has been operational). A cut set is minimal, if no component can be removed from it without losing its status as a cut set.

In terms of the structure function, a (minimal) cut set can be interpreted by a special state vector, which is known as a Minimal Cut Vector (MCV). According to the definition of a cut set, the system state for a state vector covered by a cut set is zero. Therefore, state vector $\mathbf{x}$ is a cut vector if $\phi(\mathbf{x}) = 0$.

If all components of a cut set are failed and components out of the cut set are functioning then the system is failed. A state vector that coincides with MCS is known as MCV. Using the convention that $\mathbf{y} > \mathbf{x}$, where $\mathbf{x}$ and $\mathbf{y}$ are two state vectors, for which $y_i \geq x_i$ (for $i = 1, 2, ..., n$) and there exists at least one i such that $y_i > x_i$, we say that a cut vector $\mathbf{x}$ is minimal if $\phi(\mathbf{y}) = 1$ for any $\mathbf{y} > \mathbf{x}$.

There is one-to-one correspondence between MCSs and MCVs. However, the terms MCS and MCV are slightly different. MCS is a minimal set of components, whose simultaneous failure causes the system failure, while MCV represents a situation in which the repair of any failed component results into the repair of the whole system.

The definition of MCS has been generalized for MSSs in paper [30]. The development of this conception for MCVs of MSS has been proposed in [14, 21, 22].

The generalization of the MCV definition for MSS takes into account that components of MSS have more than two states. This extension is based on the assumption that MCV is defined for every relevant system state, i.e., for states $1, 2, ..., M-1$ The definition of MCV of MSS is proposed in [14] as follows: a state vector $\mathbf{x}$ is a cut

vector for a demanded state j of the system if $\phi(\mathbf{x}) < j$. A cut vector $\mathbf{x}$ is minimal if $\phi(\mathbf{y}) \geq j$ for any $\mathbf{y} > \mathbf{x}$.

The meaning of MCVs for MSS is similar to the case of BSS, i.e., MCVs for system state j identify those situations in which the repair of any damaged component causes the improvement of the system at least to state j.

2.3 A Mathematical Method

The boundary states are one of the basic conceptions in the reliability analysis used in different mathematical methods for the computation of reliability indices and measures. For example, a boundary state is a principal item in the investigation based on Fault Trees [23, 24, 25]. These states are considered and analysed in the method of *Failure Models and Effect Analysis* (FMEA) [24, 26]. The boundary states are used in the importance analysis for the computation of the *Importance Measures* (IMs) [13, 31]. The importance analysis allows examining different aspects of reliability changes and the uncertainty in the system. IMs quantify the criticality of a particular component within the system. They have been widely used as tools for identifying system weaknesses, and to prioritise reliability improvement activities.

The most widely used IMs such as *Structural Importance* (SI), *Birnbaum Importance* (BI), and *Fussell-Vesely Importance* (FVI) are shown in Table 2 [13, 14].

Table 2 Importance measures

Short name	Description
SI	SI concentrates on the topological structure of the system and determines the proportion of working states of the system in which the working of the i-th component makes the difference between system failure and working state
BI	BI of a given component is defined as the probability that such a component is critical to MSS functioning and represents loss in MSS when the i-th component fails.
FVI	FVI quantifies the maximum decrement in MSS reliability caused by the i-th system component state deterioration and if $s = 0$, the measure allows estimating system performance level decrease for full unreliability of the i-th system component.

The Direct Partial Logic Derivative is one of possible approaches for calculation of IMs [13]. According to [13], the Direct Partial Logic Derivative has been used for calculation of SI and BI. The FVI definition is based on MCS that is calculated on the basis of the Direct Partial Logic Derivative (10) too.

In this paper, we propose to unify a method for calculation of IMs based on the Direct Partial Logic Derivatives (5), (9), and (10). These measures can be used to identify a system component that has a maximal influence on the system functioning

(system availability). In this paper, we propose to apply such an analysis for the investigation of the healthcare system.

3 Importance Analysis

The importance analysis is a part of reliability analysis that allows us to investigate the structural and topological aspects of the system from a reliability point of view. IMs permit to discover the system component with maximal/minimal influence on the system reliability/availability. This information is useful for system design and maintenance.

3.1 Structural Importance

The Structural Importance (SI) is one of the simplest measures of the component importance and this measure focuses on the topological aspects of the system. According to the definition in paper [29], this measure determines the proportion of working states of the system in which working of the i-th component makes the difference between system failure and its working:

$$IS_i^{s,j} = \frac{\rho_i^{s,j}}{\rho_i^j}, \tag{11}$$

where $\rho_i^{s,j}$ is a number of system states, when the change of the component state from s to $s-1$ results in the system reliability change from j to $j-1$; ρ_s is a number of the states for which $\phi(s_i, \mathbf{x}) = j$ and it is calculated by the structure function.

The number $\rho_i^{s,j}$ can be computed as the number of nonzero values of the Direct Partial Logic Derivative (5) .

For example, let us consider the human module of the healthcare system from paper [6] that is represented by the Fault Decision Tree (Fig. 1). Let us transform this module to MSS (Fig. 2) with the structure function defined in Table 3.

Values of SI (11) and intermediate values of $\rho_i^{s,j}$ are shown in Table 4. Therefore, the SI values in Table 4 permit to estimate the influence of every human error on the correct medical decision. The fatal influence $(IS_i^{1,1})$ on the medical mistake has doctor's errors (Fig. 1) because the values of SI are maximal for $x_3(IS_3^{1,1} = 0.1111)$, $x_4(IS_4^{1,1} = 0.1000)$, and $x_5(IS_5^{1,1} = 0.1000)$. These errors have a principal influence for some incorrections in the decision that are defined by performance level 1: $IS_3^{1,2} = 0.7647$, $IS_4^{1,2} = 0.6875$, and $IS_5^{1,2} = 0.6875$. The insignificant disadvantages (that agree with performance level 2) are caused by nurse's and doctor's errors equally.

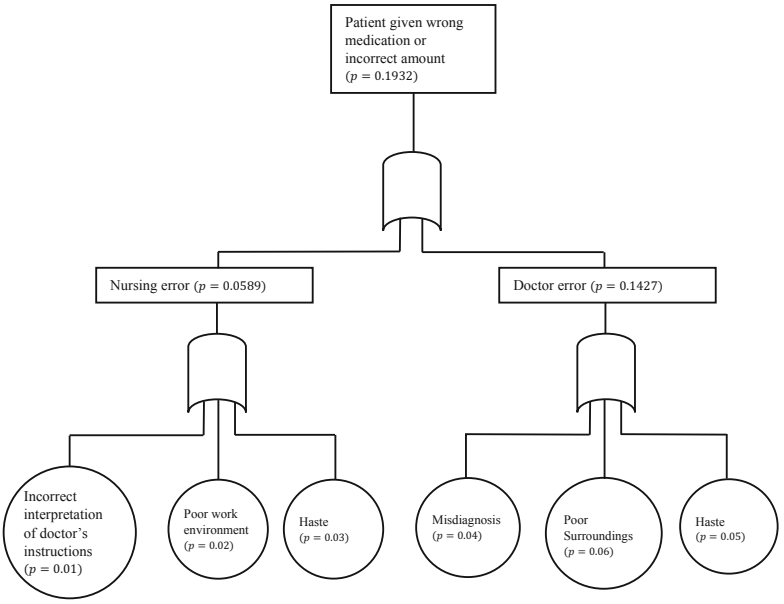
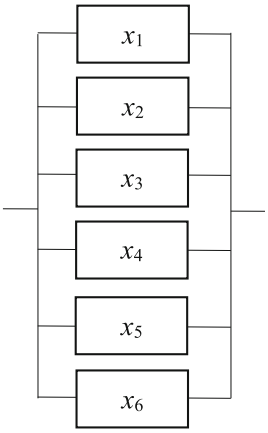


Fig. 1 The human module (from paper [6])



	$p_{0,i}$	$p_{1,i}$
Incorrect interpretation of doctor's instructions, x_1	0.01	0.99
Poor work environment, x_2	0.02	0.98
Haste, x_3	0.03	0.97
Misdiagnosis, x_4	0.04	0.96
Poor Surroundings, x_5	0.06	0.94
Haste, x_6	0.05	0.95

Fig. 2 The human module adapted to MSS from paper [6]

3.2 Birnbaum Importance

It is worth noting that SI (11) investigates correlation and influence of a system component from the point of view of the reliability/availability investigation, but this measure does not take into account the probability of a state of system components. This disadvantage is eliminated in the calculation of another importance

Table 3 The structure function of MSS for the human module in Fig. 2

	$x_1\,x_2\,x_3$							
$x_4\,x_5\,x_6$	0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1
0 0 0	0	1	1	1	1	1	1	2
0 0 1	1	1	1	1	1	1	1	2
0 1 0	1	1	1	1	1	1	1	2
0 1 1	1	2	2	2	2	2	2	3
1 0 0	1	1	1	1	1	1	2	2
1 0 1	1	2	2	2	2	2	2	3
1 1 0	1	2	2	2	2	2	2	3
1 1 1	1	2	2	3	2	3	3	3

Table 4 Structural importance for the human module in Fig. 2

i	$\rho_{s,j}$			$\rho_i^{s,j}$			IS_i		
	$\rho_{1,1}$	$\rho_{1,2}$	$\rho_{1,3}$	$\rho_i^{1,1}$	$\rho_i^{1,2}$	$\rho_i^{1,3}$	$IS_i^{1,1}$	$IS_i^{1,2}$	$IS_i^{1,3}$
1	11	15	6	1	9	5	0.0909	0.6000	0.8333
2	11	15	6	1	9	5	0.0909	0.6000	0.8333
3	12	14	6	1	7	5	0.0833	0.5000	0.8333
4	9	17	6	1	13	5	0.1111	0.7647	0.8333
5	10	16	6	1	11	5	0.1000	0.6875	0.8333
6	10	16	6	1	11	5	0.1000	0.6875	0.8333

measure that is the Birnbaum Importance (BI). BI of a given component is defined as the probability that a system is sensitive to inoperativeness of the i-th system component [14]. Let us consider the Direct Partial Logic Derivatives for calculation of BI. In paper [13], BI has been defined as:

$$IB_i^{s,j} = \Pr\left\{\frac{\partial\phi(j\to j-1)}{\partial x_i(s\to s-1)} \neq 0\right\}. \tag{12}$$

Equation (12) supposes the BI calculation as the probabilities of non-zero values of the Direct Partial Logic Derivative $\partial\phi(j\to j-1)/\partial x_i(s\to s-1)$.

For example, let us consider the system shown in Fig. 2. Probabilities of the system element reliability and unreliability are shown in this figure too. The Direct Partial Logic Derivatives for the first variable $\partial\phi(j\to j-1)/\partial x_1(1\to 0)$ are collected in Table 5.

According to data in Table 5 and equation (12) the BI measures for the first component are calculated in the following way:

Table 5 Direct Partial Logical Derivatives $\partial\phi(j \to j-1)/\partial x_1(1 \to 0)$ (10) for the human module in Fig. 2

$x_2\,x_3\,x_4\,x_5\,x_6$	$\partial\phi(1\to0)/\partial x_1(1\to0)$	$\partial\phi(2\to1)/\partial x_1(1\to0)$	$\partial\phi(3\to2)/\partial x_1(1\to0)$
0 0 0 0 0	1	0	0
0 0 0 0 1	0	0	0
0 0 0 1 0	0	0	0
0 0 0 1 1	0	1	0
0 0 1 0 0	0	0	0
0 0 1 0 1	0	1	0
0 0 1 1 0	0	1	0
0 0 1 1 1	0	1	0
0 1 0 0 0	0	0	0
0 1 0 0 1	0	0	0
0 1 0 1 0	0	0	0
0 1 0 1 1	0	0	0
0 1 1 0 0	0	0	0
0 1 1 0 1	0	0	0
0 1 1 1 0	0	0	0
0 1 1 1 1	0	0	1
1 0 0 0 0	0	0	0
1 0 0 0 1	0	0	0
1 0 0 1 0	0	0	0
1 0 0 1 1	0	0	0
1 0 1 0 0	0	1	0
1 0 1 0 1	0	0	0
1 0 1 1 0	0	0	0
1 0 1 1 1	0	0	1
1 1 0 0 0	0	1	0
1 1 0 0 1	0	1	0
1 1 0 1 0	0	1	0
1 1 0 1 1	0	0	1
1 1 1 0 0	0	1	0
1 1 1 0 1	0	0	1
1 1 1 1 0	0	0	1
1 1 1 1 1	0	0	1

$$IB_1^{1,1} = \Pr\left\{\frac{\partial\phi(1\to0)}{\partial x_1(1\to0)} = 1\right\} = p_{0,2}p_{0,3}p_{0,4}p_{0,5}p_{0,6} = 0.000000072,$$

$$IB_1^{1,2} = \Pr\left\{\frac{\partial\phi(2\to1)}{\partial x_1(1\to0)} = 1\right\} =$$
$$p_{0,2}p_{0,3}p_{0,4}p_{1,5}p_{1,6} + p_{0,2}p_{0,3}p_{1,4}p_{0,5}p_{1,6} + p_{0,2}p_{0,3}p_{1,4}p_{1,5}p_{0,6} +$$
$$p_{0,2}p_{0,3}p_{1,4}p_{1,5}p_{1,6} + p_{1,2}p_{1,3}p_{1,4}p_{0,5}p_{0,6} = 0.041442,$$

$$IB_1^{1,3} = \Pr\left\{\frac{\partial\phi(3\to2)}{\partial x_1(1\to0)} = 1\right\} =$$
$$p_{0,2}p_{1,3}p_{1,4}p_{1,5}p_{1,6} + p_{1,2}p_{0,3}p_{1,4}p_{1,5}p_{1,6} + p_{1,2}p_{1,3}p_{0,4}p_{1,5}p_{1,6} +$$
$$p_{1,2}p_{1,3}p_{1,4}p_{0,5}p_{1,6} + p_{1,2}p_{1,3}p_{1,4}p_{1,5}p_{0,6} = 0.985629.$$

The BI measures for all system components are collected in Table 6. These measures are calculated according to (12) and on the basis of the Direct Partial Logic Derivatives for the structure function of the system (the human module) in Fig. 2.

Table 6 BI for the human module in Fig. 2

i	$IB_i^{1,1}$	$IB_i^{1,2}$	$IB_i^{1,3}$
1	0,000000072	0,007487	0,170699
2	0,000000036	0,007259	0,163955
3	0,000000024	0,004351	0,156896
4	0,000000018	0,006213	0,149505
5	0,000000012	0,003667	0,133664
6	0,000000014	0,003931	0,141767

Therefore, the fatal result caused by the doctor's or nurse's errors has minimal probability. In this situation the "*Incorrect interpretation of doctor's instructions*" has a maximal influence on the fatal result according to input data (Fig. 1 and Fig. 2). This error, according to the structure function (Table 1) and probabilities of the module components (Fig. 1), has a principal influence on other performance levels too. BI is a more informative measure because it considers the structure function and the probabilities of component states.

3.3 Fussell-Vesely Importance

There is one more measure that allows us to investigate an influence of the component state change on the system performance level (reliability/availability). It is the Fussell-Vesely Importance (FVI) that represents the contribution of each component to the system failure probability and, for BSS, it is calculated by the equation [27]:

$$IFV_i = \frac{F_{min\_cut}(x_i)}{Q}. \tag{13}$$

where $F_{min\_cut}(x_i)$ is the system minimal cut that includes the i-th system component and Q is the function of the system unreliability [27]:

$$Q = \Pr\{\phi(\mathbf{x}) = 0\}. \tag{14}$$

As we can see, the calculation of this measure is based on minimal cut sets, that have been considered above in Section 2.

The FVI measure for MSS is defined by the conception of MCVs that are defined and computed according to (10). FVI for MSS is defined as [13]:

$$IFV_i = \frac{\Pr\left\{MCV_i^j\right\}}{1-R(j)}, \tag{15}$$

where MCV_i^j is MCV for system performance level j that includes the i-th system component; $R(j)$ is the reliability for system performance level j [14]:

$$R(j) = \Pr\{\phi(\mathbf{x}) \le j\}; U(j) = 1 - R(j) = \Pr\{\phi(\mathbf{x}) < j\}. \tag{16}$$

MCV_i^j is MCV for system performance level j that includes the i-th system component and $\Pr\{MCV_i^j\}$ is the probability that system components are in such states that their improvement results in a system repair to performance level j or greater, i.e.:

$$\Pr\left\{MCV_i^j\right\} = \Pr\left\{\mathbf{x} \le MCV_{i,1}^j \text{ OR } \mathbf{x} \le MCV_{i,2}^j \text{ OR ... OR } \mathbf{x} \le MCV_{i,l_i^j}^j\right\}, \tag{17}$$

where l_i^j is the number of MCVs for system performance level j that include component i and $MCV_{i,t}^j$ is the t-th MCV from the set MCV_i^j.

In the human module in Fig. 2, the reliability of every component, i.e., probability that the system component is in state 1, is high and therefore, Expression (17) can be estimated as follows [28]:

$$\Pr\left\{MCV_i^j\right\} \approx \sum_{t=1}^{l_i^j} \Pr\left\{\mathbf{x} \le MCV_{i,t}^j\right\}. \tag{18}$$

The FVI measure for the human module can be computed as follows:

$$IFV_i^j \approx \frac{1}{1-R(j)} \sum_{t=1}^{l_i^j} \Pr\left\{\mathbf{x} \le MCV_{i,t}^j\right\}. \tag{19}$$

Now, let us compute the FVI measures for individual components of the human module in Fig. 2. MCVs of the human module have to be computed firstly. In Section 2, it was mentioned that Direct Partial Logic Derivatives can be used for this task.

Let us assume that we want to compute MCVs for performance level 2 of the human module in Fig. 2. Firstly, the merge (10) of Direct Partial Logic Derivatives unions (9) and next their intersection have to be calculated. In Table 7, the merge of Direct Partial Logic Derivatives according to components 1, 2, and 3 (the white columns) and their intersection (the gray columns) for state 2 of the human module are computed. This intersection identifies situations in which the repair of any failed component from the set of components (1, 2, and 3) causes the transition of the system from the performance level less than 2 to level 2 or greater. However, MCVs for this performance levels are defined as state vectors in which the repair of any failed system component results in this type of a system improvement. Therefore,

the merge of Direct Partial Logic Derivatives according to other components, i.e., 4, 5, and 6 and their intersection have to be calculated too (Table 8).

Table 7 The merge (10) of Direct Partial Logic Derivatives unions (9) for state 2 of the human module in Fig. 2 according to variables x_1, x_2, and x_3 and their intersection

	$x_1\,x_2\,x_3$															
$x_4\,x_5\,x_6$	0 0 0		0 0 1		0 1 0		0 1 1		1 0 0		1 0 1		1 1 0		1 1 1	
0 0 0	0 0 0	0	0 0 *	0	0 * 0	0	1 * *	1	* 0 0	0	* 1 *	1	* * 1	1	* * *	*
0 0 1	0 0 0	0	0 0 *	0	0 * 0	0	1 * *	1	* 0 0	0	* 1 *	1	* * 1	1	* * *	*
0 1 0	0 0 0	0	0 0 *	0	0 * 0	0	1 * *	1	* 0 0	0	* 1 *	1	* * 1	1	* * *	*
0 1 1	1 1 1	1	0 0 *	0	0 * 0	0	0 * *	0	* 0 0	0	* 0 *	0	* * 0	0	* * *	*
1 0 0	0 0 0	0	0 0 *	0	1 * 0	0	1 * *	1	* 1 0	0	* 1 *	1	* * 0	0	* * *	*
1 0 1	1 1 1	1	0 0 *	0	0 * 0	0	0 * *	0	* 0 0	0	* 0 *	0	* * 0	0	* * *	*
1 1 0	1 1 1	1	0 0 *	0	0 * 0	0	0 * *	0	* 0 0	0	* 0 *	0	* * 0	0	* * *	*
1 1 1	1 1 1	1	0 0 *	0	0 * 0	0	0 * *	0	* 0 0	0	* 0 *	0	* * 0	0	* * *	*

Table 8 The merge (10) of Direct Partial Logic Derivatives unions (9) for state 2 of the human module in Fig. 2 according to variables x_4, x_5, and x_6 and their intersection

	$x_1\,x_2\,x_3$															
$x_4\,x_5\,x_6$	0 0 0		0 0 1		0 1 0		0 1 1		1 0 0		1 0 1		1 1 0		1 1 1	
0 0 0	0 0 0	0	0 0 0	0	0 0 0	0	0 0 0	0	0 0 0	0	0 0 0	0	1 0 0	0	0 0 0	0
0 0 1	0 0 *	0	1 1 *	1	1 1 *	1	1 1 *	1	1 1 *	1	1 1 *	1	1 1 *	1	0 0 *	0
0 1 0	0 * 0	0	1 * 1	1	1 * 1	1	1 * 1	1	1 * 1	1	1 * 1	1	1 * 1	1	0 * 0	0
0 1 1	0 * *	0	0 * *	0	0 * *	0	0 * *	0	0 * *	0	0 * *	0	0 * *	0	0 * *	0
1 0 0	* 0 0	0	* 1 1	1	* 1 1	1	* 1 1	1	* 1 1	1	* 1 1	1	* 0 0	0	* 0 0	0
1 0 1	* 0 *	0	* 0 *	0	* 0 *	0	* 0 *	0	* 0 *	0	* 0 *	0	* 0 *	0	* 0 *	0
1 1 0	* * 0	0	* * 0	0	* * 0	0	* * 0	0	* * 0	0	* * 0	0	* * 0	0	* * 0	0
1 1 1	* * *	*	* * *	*	* * *	*	* * *	*	* * *	*	* * *	*	* * *	*	* * *	*

In Table 9, there is the computed intersection (the gray columns) of intersections (the white columns) from Table 7 and Table 8. This intersection identifies all situations in which the repair of any failed component coincides with the transition of the human module from the performance level less than 2 to level 2 or greater, i.e., it reveals MCVs of the human module for performance level 2.

The previous process can be performed for other performance levels of the system in Fig. 2 to find MCVs for every relevant performance level of the system, i.e., for levels 1, 2, and 3.

When MCV contains value 0 at position i, then it is MCV for the i-th component. Using this rule, MCVs for a given performance level of the human module can be

Table 9 The intersections from Table 7 and Table 8

	$x_1\,x_2\,x_3$															
$x_4\,x_5\,x_6$	0 0 0		0 0 1		0 1 0		0 1 1		1 0 0		1 0 1		1 1 0		1 1 1	
0 0 0	0 0	0	0 0	0	0 0	0	1 0	0	0 0	0	1 0	0	1 0	0	* 0	0
0 0 1	0 0	0	0 1	0	0 1	0	1 1	**1**	0 1	0	1 1	**1**	1 1	**1**	* 0	0
0 1 0	0 0	0	0 1	0	0 1	0	1 1	**1**	0 1	0	1 1	**1**	1 1	**1**	* 0	0
0 1 1	1 0	0	0 0	0	0 0	0	0 0	0	0 0	0	0 0	0	0 0	0	* 0	0
1 0 0	0 0	0	0 1	0	0 1	0	1 1	**1**	0 1	0	1 1	**1**	0 0	0	* 0	0
1 0 1	1 0	0	0 0	0	0 0	0	0 0	0	0 0	0	0 0	0	0 0	0	* 0	0
1 1 0	1 0	0	0 0	0	0 0	0	0 0	0	0 0	0	0 0	0	0 0	0	* 0	0
1 1 1	1 *	**1**	0 *	0	0 *	0	0 *	0	0 *	0	0 *	0	0 *	0	* *	*

split into six groups (one group for one component) and the i-th group contains only those MCVs that have value 0 at position i, i.e., the i-th group represents MCVs for the i-th component (Table 10).

Table 10 MCVs for components and states of the human module in Fig. 2

System state	Component					
	1	2	3	4	5	6
1	(0,0,0,0,0,0)	(0,0,0,0,0,0)	(0,0,0,0,0,0)	(0,0,0,0,0,0)	(0,0,0,0,0,0)	(0,0,0,0,0,0)
2	(0,0,0,1,1,1) (0,1,1,0,0,1) (0,1,1,0,1,0) (0,1,1,1,0,0)	(0,0,0,1,1,1) (1,0,1,0,0,1) (1,0,1,0,1,0) (1,0,1,1,0,0)	(0,0,0,1,1,1) (1,1,0,0,0,1) (1,1,0,0,1,0)	(0,1,1,0,0,1) (0,1,1,0,1,0) (1,0,1,0,0,1) (1,0,1,0,1,0) (1,1,0,0,0,1) (1,1,0,0,1,0)	(0,1,1,0,0,1) (0,1,1,1,0,0) (1,0,1,0,0,1) (1,0,1,1,0,0) (1,1,0,0,0,1)	(0,1,1,0,1,0) (0,1,1,1,0,0) (1,0,1,0,1,0) (1,0,1,1,0,0) (1,1,0,0,1,0)
3	(0,0,1,1,1,1) (0,1,0,1,1,1) (0,1,1,0,1,1) (0,1,1,1,0,1) (0,1,1,1,1,0)	(0,0,1,1,1,1) (1,0,0,1,1,1) (1,0,1,0,1,1) (1,0,1,1,0,1) (1,0,1,1,1,0)	(0,1,0,1,1,1) (1,0,0,1,1,1) (1,1,0,0,1,1) (1,1,0,1,0,1) (1,1,0,1,1,0)	(0,1,1,0,1,1) (1,0,1,0,1,1) (1,1,0,0,1,1) (1,1,1,0,0,1) (1,1,1,0,1,0)	(0,1,1,1,0,1) (1,0,1,1,0,1) (1,1,0,1,0,1) (1,1,1,0,0,1) (1,1,1,1,0,0)	(0,1,1,1,1,0) (1,0,1,1,1,0) (1,1,0,1,1,0) (1,1,1,0,1,0) (1,1,1,1,0,0)

In the next step, the probabilities $\Pr\left\{MCV_i^j\right\}$ for the i-th component and the j-th system performance level can be computed using (18). These probabilities are calculated in Table 11 and they represent the contribution of the i-th component failure to the system unreliability (16).

Finally, when the FVI measures of individual components of the human module should be calculated, the unreliabilities for individual system performance levels have to be computed (Table 12).

Using data from Table 11 and Table 12, the FVI measures for individual components and system performance levels of the human module in Fig. 2 are computed in Table 13. According to this table, the probability that the failure of any system

Table 11 Contributions of components of the human module in Fig. 2 to the module unreliability

Component (i)	$\Pr\{MCV_i^1\}$	$\Pr\{MCV_i^2\}$	$\Pr\{MCV_i^3\}$
1	0.00000000072	0.000080	0.0020
2	0.00000000072	0.000154	0.0038
3	0.00000000072	0.000138	0.0054
4	0.00000000072	0.000264	0.0068
5	0.00000000072	0.000234	0.0090
6	0.00000000072	0.000210	0.0080

Table 12 Reliabilities and unreliabilities of states of the human module in Fig. 2

System state (j)	$R(j)$	$U(j) = 1 - R(j)$
1	0.99999999928	0.00000000072
2	0.99965610128	0.00034389872
3	0.98392198200	0.01607801800

component contributes to the total failure of the human module, i.e., IFV_i^1, is the same for all components. This is caused by the fact that there is only one MCV for system performance level 1 and it is shared by all system components.

The values of IFV_i^2 describe the probabilities that the failure of the i-th component contributes to human module unreliability (16) with respect to performance level 2 of the system. According to these values, the "*Misdiagnosis*" has the greatest and the "*Incorrect interpretation of doctor's instructions*" have the smallest contribution.

In the last column of Table 13, there are values of FVI for performance level 3 and they indicate that the "*Poor Surroundings*" have the greatest contribution to the system unreliability (16) defined with respect to performance level 3 of the human module and the "*Incorrect interpretation of doctor's instructions*" have the smallest contribution.

Table 13 The FVIs for the human module in Fig. 2

Component (i)	IFV_i^1	IFV_i^2	IFV_i^3
1	1	0.232627	0.124393
2	1	0.447806	0.236348
3	1	0.401281	0.335862
4	1	0.767668	0.422938
5	1	0.680433	0.559770
6	1	0.610645	0.497574

As we can see from Table 6 and Table 13, BI and FVI give the different order of influence of individual system components on the reliability of the human module in Fig. 2. Mainly, this is caused by the fact that BI does not take into account the

reliability of a component for which it is computed. The "*Incorrect interpretation of doctor's instructions*" have the greatest consequence on the human module reliability (it is defined by BI), but this event is very rare and therefore, the probability, that this event will contribute to the system unreliability (it is defined by FVI), is very small.

4 Discussion and Conclusions

The analysis of the medical mistake, occurring in the last decades, has clearly shown that the organizational and human factor play a significant role in the risk of misdiagnosis and wrong treatment [8, 11]. This is also due to the fact that the reliability of the technical components has been significantly improved in recent years. As a consequence, the influence of the errors of the organization managing and human operators on system operation has been significantly increased. Therefore, the correct reliability analysis of a healthcare system needs to be considered in relation to three basic components of the system [9]: technical (hardware and software), human and organizational. The analysis of these components needs similar algorithms, but at present, the reliability estimation of these components is based on the different approaches investigated by different methods. In this paper, we have developed one of possible ways for the reliability analysis of a healthcare system based on the interpretation of a healthcare system as MSS. A new approach for the quantification of a healthcare system as MSS is developed. The background of this analysis is the importance analysis that allows us to estimate an influence of functioning or failure of different system components on the system reliability (availability, performance) change.

In this paper, new algorithms for calculation of IMs (Table 2) for MSS analysis are considered. These algorithms are implemented on the basis of the mathematical approach - Multiple-Valued Logic, especially Direct Partial Logic Derivatives. Application of these measures in the analysis of the healthcare system reliability is presented. In this paper, these measures and algorithms for the calculation are used for the human module of the healthcare system that is not a typical application of methods of the importance analysis. The proposed algorithms can be used for other components of a healthcare system. In this case, the investigated system includes technical components (hardware and software), human factor and organizational components, and it is interpreted as MSS. This interpretation of a system allows us to investigate some performance levels of system functioning that do not include the level of functioning and fault only. Therefore, the proposed algorithms can be used to investigate healthcare components (for example, as technical or human) and the reliability behaviour of a healthcare system at large.

Acknowledgements. This work was supported by the grant of 7th RTD Framework Programme No 610425 (RASimAs), the grant Vega 1/0498/14, and the grant of Slovak Research and Development Agency SKPL0023-12.

References

1. Taylor, E.F.: The reliability engineer in the health care system. In: Proc IEEE the 18th Annual Reliability & Maintainability Symposium, USA, pp. 245–248 (1972)
2. Cohen, T.: Medical and Information Technologies Converge. IEEE Engineering in Medicine and Biology Magazine 23(3), 59–65 (2004)
3. Dhillon, B.S.: Medical Device Reliability and Associated Areas. CRC Press (2000)
4. Taleb-Bendiab, A., England, D., Randles, M., Miseldine, P., Murphy, K.: A principled approach to the design of healthcare systems: Autonomy vs. governance. Reliability Engineering and System Safety 91(12), 1576–1585 (2006)
5. Deeter, J., Rantanen, E.: Human Reliability Analysis in Healthcare. In: Proc. of 2012 Symposium on Human Factors and Ergonomics in Health Care, Baltimore, USA, pp. 45–51 (2012)
6. Lyons, M., Adams, S., Woloshynowych, M., Vincent, C.: Human reliability analysis in healthcare: A review of techniques. Int. Journal of Risk & Safety in Medicine. 16(4), 223–237 (2004)
7. Spyrou, S., Bamidis, P.D., Maglaveras, N., Pangalos, G., Pappas, C.: A methodology for reliability analysis in health networks. IEEE Trans. on Information Technology in Biomedicine 12(3), 377–386 (2008)
8. Bogner, M.: Human Error in Medicine. Lawrence Erlabaum Associates, Hillsdale (1994)
9. Zaitseva, E., Levashenko, V., Rusin, M.: Reliability Analysis of Healthcare System. In: Proc. of the IEEE Federated Conf. on Computer Science & Information Systems, Szczecin, Poland, pp. 169–175 (2011)
10. Zio, E.: Reliability engineering: Old problems and new challenges Importance analysis by logical differential calculus. Reliability Engineering and Syst. Safety. 141, 9125–9141 (2009)
11. Dhillon, B.S.: Human Reliability and Error in Medical System. World Scientific (2003)
12. Zaitseva, E., Kostolny, J., Kvassay, M., Levashenko, V., Pancerz, K.: Failure Analysis and Estimation of the Healthcare System. In: Proc. of the IEEE Federated Conf. on Computer Science & Information Systems, Krakow, Poland, pp. 235–240 (2013)
13. Zaitseva, E.: Importance analysis of a multi-state system based on multiple-valued logic methods. In: Recent Advances in System Reliability: Signatures. Multi-state Systems and Statistical Inference, pp. 113–134. Springer, London (2012)
14. Lisnianski, A., Levitin, G.: Multi-State System Reliability. Assessment, Optimization and Applications. World Scientific (2003)
15. Natvig, B.: Multistate systems reliability theory with applications. Wiley, New York (2011)
16. Li, Y., Zio, E.: A Multi-State Model for the Reliability Assessment of a Distributed Generation System via Universal Generating Function. Reliability Engineering and System Safety 106(1), 28–36 (2012)
17. Zaitseva, E., Levashenko, V.: Importance analysis by logical differential calculus. Automation and Remote Control 74(2), 171–182 (2013)
18. Shooman, M.L.: Probabilistic Reliability: An Engineering Approach. McGraw-Hill, New York (1968)
19. Al-Muhaini, M., Heydt, G.T.: Minimal cut sets, Petri nets, and prime number encoding in distribution system reliability evaluation. In: Proc. Transmission and Distribution Conference and Exposition, Orlando, USA, pp. 1–8 (2012)
20. Tapia, M.A., Guima, T.A., Katbab, A.: Calculus for a multivalued-logic algebraic system. Applied Mathematics and Computation 42, 255–285 (1991)

21. Soh, S., Rai, S.: An efficient cutset approach for evaluating communication-network reliability with heterogeneous link-capacities. IEEE Trans. on Reliability 54(1), 133–144 (2005)
22. Yeh, W.: A fast algorithm for searching all multi-state minimal cuts. IEEE Transactions on Reliability 57(4), 581–588 (2008)
23. Contini, S., Matuzas, V.: New methods to determine the importance measures of initiating and enabling events in fault tree analysis. Reliability Engineering & System Safety 96(7), 775–784 (2011)
24. Farcasiu, M., Prisecaru, I.: MMOSA A new approach of the human and organizational factor analysis in PSA. Reliability Engineering and System Safety 123(1), 91–98 (2014)
25. Merle, G., Roussel, J.M., Lesage, J.J.: Algebraic determination of the structure function of Dynamic Fault Trees. Reliability Engineering and System Safety 96(2), 267–277 (2011)
26. Seyed-Hosseini, S.M., Safaei, N., Asgharpour, M.J.: Reprioritization of failures in a system failure mode and effects analysis by decision making trial and evaluation laboratory technique. Reliability Engineering & System Safety 91(8), 872–881 (2006)
27. Meng, F.C.: On some structural importance of system components. Journal of Data Science 7, 277–283 (2009)
28. Zio, E.: An Intriduction to the Basics of reliability and Risk Analysis. World Scientific (2007)
29. Armstrong, M.: Reliability-importance and dual failure-mode components. IEEE Trans. on Reliability. 46(2), 212–221 (1997)
30. Butler, D.A.: A complete importance ranking for components of binary coherent systems with extensions to multi-state systems. Naval Research Logistics 4, 565–578 (1979)
31. Vaurio, J.K.: Ideas and developments in importance measures and fault-tree techniques for reliability and risk analysis. Reliability Engineering & System Safety 95(2), 99–107 (2010)

Application of Selected Classification Methods for Detection of Hashimoto's Thyroiditis on the Basis of Ultrasound Images

Zbigniew Omiotek, Andrzej Burda, and Waldemar Wójcik

Abstract. The article presents methods for classification of ultrasound thyroid images. These methods allow us to classify analyzed cases as sick or healthy. They allow us to separate a fairly large group of incorrectly classified cases as well. According to the authors, this group may contain features of the early stage of Hashimoto's disease. Decision tree induction and artificial neural networks have been used to build classification models. Test results showed that the proposed methods can provide a starting point to build a support system in the process of medical diagnosis.

1 Introduction

The thyroid is an endocrine gland located in the bottom front part of the neck (Figure 1). It is composed of two lateral lobes connected by a narrow passage. Some people have the third piece, called a pyramidal lobe. The thyroid gland produces hormones called triiodothyronine (T3), thyroxine (T4) and calcitonin. They affect the metabolism and the calcium-phosphate.

One of the most common diseases of this gland is Hashimoto's disease. In addition to its classical form, there are several varieties of this disease: postpartum

Zbigniew Omiotek
Electrical Engineering and Computer Science Faculty, Lublin University of Technology, Lublin, Poland
e-mail: zmiotek@gmail.com

Andrzej Burda
Department of Computer Science and Knowledge Engineering, University of Management and Administration, Zamość, Poland
e-mail: aburda@wszia.edu.pl

Waldemar Wójcik
Electrical Engineering and Computer Science Faculty, Lublin University of Technology, Lublin, Poland
e-mail: waldemar.wojcik@pollub.pl

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*, Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_2

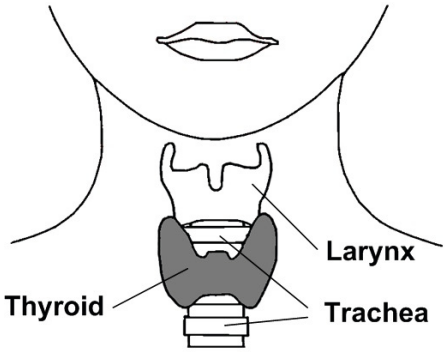


Fig. 1 The location of the thyroid gland

thyroiditis, asymptomatic thyroiditis, atrophic autoimmune thyroiditis, focal lymphocytic thyroiditis. The disease is named after the Japanese physician - Hakuru Hashimoto. This physician discovered and diagnosed four cases of the disease, and introduced its classic form in 1912. Hashimoto's disease is an autoimmune disease - associated with impaired immune systems. It is also often called a chronic lymphocytic inflammation of the thyroid gland. This disease is the most common cause of hypothyroidism, and the most common type of inflammation of the gland. It occurs in approximately 2% of the population, and it is estimated that women suffer from it more than 10 times more often than men. It is usually diagnosed in the period between 30th and 50th year of age.

Hashimoto's disease is manifested in the appearance of the characteristic inflammatory changes within the thyroid and it can cause both hyperplasia and atrophy of the gland. It may take the form of focal and diffuse. The focal form of the disease causes moderate changes in the thyroid and usually does not cause its destruction. The diffuse form of the disease affects the entire thyroid gland or a substantial part of it. The thyroid gland with inflammation expands and forms a symmetrical, hard goiter of more or less smooth surface. The treatment of the diffuse form of the disease is generally ineffective [9].

The diagnosis of early stages of Hashimoto's disease is very important for the effective treatment. Patients with such a diagnosis may be much earlier referred for a specialized laboratory and imaging testing. The use of proper treatment procedures, as early as possible, is very important for ensuring the effectiveness of therapy. Despite the development of methods for diagnosing Hashimoto's disease, the accurate diagnosis of its early stage is still a big problem. This follows from the fact that symptoms of thyroid disorders are highly variable and they have often slow dynamics of their growth. The disease is usually diagnosed in the advanced stage when there already occurs hypothyroidism and the appropriate treatment is difficult or impossible. Therefore, an extremely important issue is to increase the effectiveness of existing diagnostic methods, so that diagnosis can be made as early as possible.

Many expert systems have been developed so far, which detect a healthy thyroid gland and the cases of hypothyroidism and hyperthyroidism with high accuracy. They use different machine learning methods. The recent research in this area indicate that these systems are mostly based on hormonal tests. For instance, they use the information about the level of thyrotropin (TSH) produced by the pituitary gland and triiodothyronine (T3) and thyroxine (T4), which are produced in the thyroid [6], [7], [8], [16], [20], [30], [32].

The basic method for diagnosis of the thyroid is still an ultrasonography. Therefore, the possibility of using ultrasound images to aid the process of this diagnosis is currently an area of active research. Many methods for image processing and analysis have been developed, so that the information obtained from ultrasound images can be fuller and more accurate. In [17], a method for automatic measurement of diameter of the thyroid gland was presented. It gives the information about the size of patient's goiter. In [28], an algorithm for improving the contrast of an ultrasound image was presented. It allows us to perform more precise segmentation of the thyroid area by eliminating the stains on the image. New algorithms for determining the boundaries of the thyroid gland and extraction of image features are proposed. They are resistant to noise and accurately describe the echogenic properties of the analyzed tissue [19], [18]. Particular emphasis is placed on detection and identification of thyroid nodules, which are the symptoms of various pathological conditions of the gland. Various machine learning methods are used for this purpose.

The use of genetic algorithms for segmentation of the nodules is proposed in [14]. The presented method allowed reducing the negative effects of non-uniform intensity of the ultrasound image. Classifiers based on the support vector machine are very popular as well [4], [5], [15]. Good results can also be achieved by combining the possibilities of several classifiers. The project of a biometric system that uses unique features extracted from ultrasound images of the thyroid was presented in [27]. In this system, the combination of classifiers (k-nearest neighbors, maximum a posteriori and entropy distance) was used. The proposed method can also be successfully used in a diagnostic tool.

There is a relatively small number of publications presenting the results of studies aimed at identification of Hashimoto's disease on the basis of ultrasound images. The influence of the size of image samples on the reproducibility of classification, depending on changes in the scanning parameters and the segmentation of the thyroid, was analyzed in [29]. An interesting approach to the automatic evaluation of echogenicity of the gland parenchyma was presented in [21]. In this article, a method, that uses the information about the texture of the artery located in the thyroid as a reference point, was presented. On this basis, regions of interest are detected in tested tissue. k-NN classifiers are very often used to identify a healthy thyroid and the cases diagnosed with Hashimoto's disease. Such an approach was presented in [2]. In this article, the advantages of the wavelet transformation as a feature vector extraction method for discrimination purpose were shown as well. Good results in detection of the parenchymal inflammation caused by the Hashimoto's disease have been achieved by using the naive Bayesian classifiers [26]. This type of classifiers allows us to achieve the detection accuracy at the level of 92%.

A review of the latest publications leads to the conclusion that the current research focuses on identifying two typical states of the thyroid based on the ultrasound images. The first one is the healthy gland, and the second one is the advanced Hashimoto's disease. However, the authors do not know any research on detection of early stage of the disease mentioned above. This fact justifies the need for research in this area which seems to be particularly important for the effectiveness of the therapy.

An ultrasonography is based on the phenomenon of propagation, scattering and reflection of the ultrasonic wave on the boundary of two tissues, assuming a constant wave speed in different tissues. Different tissues reflect the ultrasound in a different way, therefore it is possible to get the image of the thyroid crosssection on the monitor. Ultrasound images allow us to assess the size of the thyroid, echogenicity and heterogeneity of parenchyma, the presence of nodules, calcifications and fluids (cysts). A typical ultrasound image for Hashimoto's disease has reduced echogenicity and heterogeneity of the parenchyma (Figure 2).

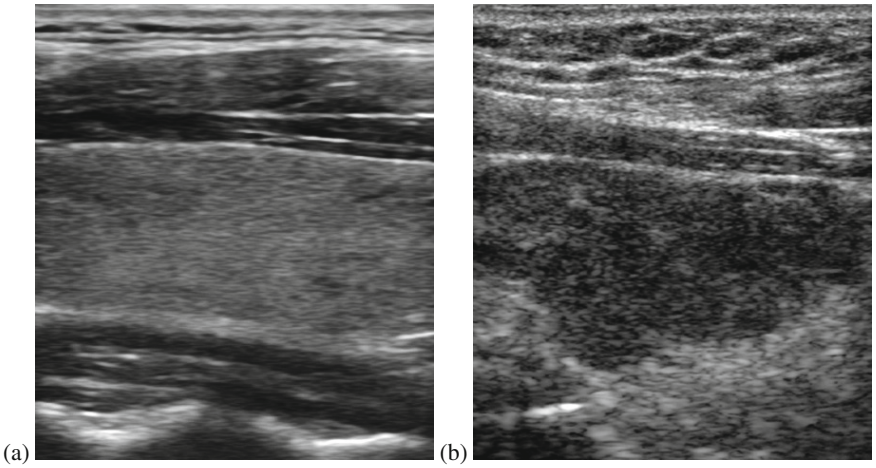


Fig. 2 Ultrasound thyroid images: (a) healthy gland (homogeneous parenchyma, normal echogenicity), (b) gland with a diagnosis of Hashimoto (heterogeneous parenchyma, hypoechogenicity)

A physician may relatively easy diagnose the cases of healthy thyroid and the gland affected by the advanced form of the disease on the basis of ultrasound images. However, such a diagnosis method for the initial stage of Hashimoto's disease is not an easy task. Therefore, it seems to be advisable to develop a computer method for an automatic identification and classification of ultrasound thyroid images. This method would be able to help the physician in the diagnosis of cases affected by the early stage of the disease. In this work, the concepts of such methods, which use decision tree induction and artificial neural networks in the classification process,

have been proposed. They allow us to classify the analyzed cases into one of two categories: sick or healthy.

This paper is organized in five sections. Section 2 presents the methods for classification using decision tree induction and a multilayer perceptron neural network. In Section 3, the processing and analysis of USG images used for the test have been described. The procedures for building and testing the classifier models have been presented in Section 4. Finally, conclusions are included in Section 5.

2 Object Classification Methods

Classification is one of the oldest and most important methods for data mining. It provides the mapping of data into a set of predefined classes. The model is built on the basis of a database content. It is used for classifying new objects in a database or for better understanding of existing distribution of objects into the predefined classes. Classification has found many applications, e.g. an automatic detection of objects in large image databases, decision support in bank loans, identifying trends in financial markets. It is also used in medical systems. The rules for the classification of various diseases in the medical database may be found. Next, these rules can be used for an automatic diagnosis of new patients. Classification is a two step process. At the first stage, the model (classifier) is constructed, which describes a predefined set of data classes or a set of concepts. The input data is the training set of samples (cases, observations) which is the list of values of descriptive attributes and the selected decision attribute. On the basis of descriptive values, the classifier learns how to assign a proper value of a decision attribute to each case in the training set, so that the classification error for this set is the smallest one. In the second stage, the resulting model is used for classification (prediction) of new objects.

In the literature, many methods for classification are proposed. These are, e.g. classification by decision tree induction, artificial neural networks, rough sets, Bayes classifiers, metaheuristics (e.g. genetic algorithms), k-nearest neighbors methods, statistical analysis and many other methods that are still under development. Among these methods for classification, the most popular is decision tree induction. It is particularly attractive because of the intuitive, understandable for a human, and therefore friendly, way of knowledge representation [23], [24], [31].

2.1 Classification by Decision Tree Induction

A decision tree can be constructed relatively quickly as compared to other methods for classification. The advantage of its use is also the scalability for large data sets and the ability to use multidimensional data. In addition, the accuracy of this method is comparable to the accuracy of other methods for classification. Most available commercial tools for data mining are based on a decision tree model. On the other

hand, the main disadvantage of this method in its basic form, is high sensitivity to missing values of attributes, because it is based on the assumption that the information stored in the database of cases is fully available. Its disadvantage is also the lack of possibility to evaluate the correlation between attributes.

In the method based on decision tree induction, the classification result is a directed acyclic graph, which has the structure of a tree. Each vertex (node) of the graph represents the test for attributes, each arc represents the result of the test, and each leaf represents a single class or the class values distribution. The highest node is called the root of the tree. A decision tree splits recursively the training set on partitions until each partition contains data belonging to one class or the partition is dominated by data belonging to one class. It should be noted that the partition size is limited. Each internal vertex of the tree, the so-called split point, performs a test on the attributes and in this way it divides the data set into partitions.

The basic algorithm for decision tree construction is a greedy algorithm. It creates a tree recursively with a top-down technique in a way of ”divide and conquer”. There are many variants of this algorithm but the most commonly used are ID3 and C4.5 algorithms. Other methods are CART method, which is derived from statistics and CHAID, which is associated with image recognition. These algorithms differ mainly in the criterion of division, i.e., the way of creating new nodes inside the tree. Classification criterion should maximize the accuracy of the constructed tree, or, in other words, minimize misclassification of data. A decision tree is usually constructed in two phases. In the first phase, called a construction phase, growth, or induction, a tree is built on the basis of a training database. In the second phase, called a pruning or reduction phase, the tree is pruning in order to improve the accuracy and to keep the independence from the effects of overtraining. In the pruning phase, the branches representing the singularities and noise existing in the data are identified and removed. Two main strategies are related to the pruning of a tree. These strategies are the postpruning, where the complete decision tree is constructed and its unreliable parts are removed, and prepruning, where the branch cases are developed, if the information starts to be unreliable. The postpruning is preferred in practice, because the prepruning often causes the effect called ”early stop”.

2.2 Classification by Multilayer Perceptron Neural Network

The most popular type of an artificial neural network to solve the classification problems is a multilayer perceptron (MLP) [33]. Due to the typical use of the hyperbolic tangent function and the logistic function as the activation functions in this model, a separation of cases on certain categories runs along the hyperplanes determined in the process of network learning. MLP learning process is usually based on the analytical method for determining the gradient of the classification error function. Afterwards, the network parameters are modified in the direction of the greatest decrease in the value of this function. This method of learning is known as a backward error propagation algorithm.

It should be emphasized that the purpose of the network training is to ensure the correct classification of objects outside of a training set. We then say that the network generalizes of knowledge contained in a training set on the basis of the cases presented in the learning process. Otherwise, if the classification results on a training set are not confirmed on an independent testing set, we deal with the syndrome of ”learning by heart”.

To avoid this phenomenon, the learning process (matching of the discrimination hyperplane) should be stopped at the right time. A special validation set is used for this purpose. If the classification error for this set of cases, which do not take direct part in the learning process, reaches its minimum, then we recognize that we deal with an optimal model.

We find a deficiency of the learning cases in most of the real classification tasks. To limit the risk of a network overtraining, the number of attributes in the network input is reduced. The simplest, and the most common accepted rule is an estimation of this number according to the empirical formula:

$$L = 2^N, \tag{1}$$

where:

- L is the number of cases in the training set,
- N is the number of neurons in the network input.

3 Image Processing and Analysis

The study used a series of thyroid ultrasound images belonging to 66 patients. There were 31 healthy patients and 35 patients with a diagnosis of Hashimoto’s disease. The images have been extracted from one to three regions of interest (ROI) of size 200×150 pixels for each series. These ROIs covered the central areas of the left and right gland of the thyroid. As a result, 54 samples for the cases diagnosed as healthy, and 85 for sick cases (139 cases in total) have been obtained. Afterwards, extracted fragments of images were normalized with a linear correction in Matlab. An example of a sample image, before and after normalization, has been shown in Figure 3.

The MaZda computer program has been used to extract discriminant features of analyzed images. It was developed in the Institute of Electronics at the Lodz University of Technology and made available for research on the Internet, free of charge. MaZda allows the analysis of the gray texture images and calculates an extensive set of 283 discriminant features. Detailed description of these features can be found in [10], [11], [13], [22], and in documentation of MaZda.

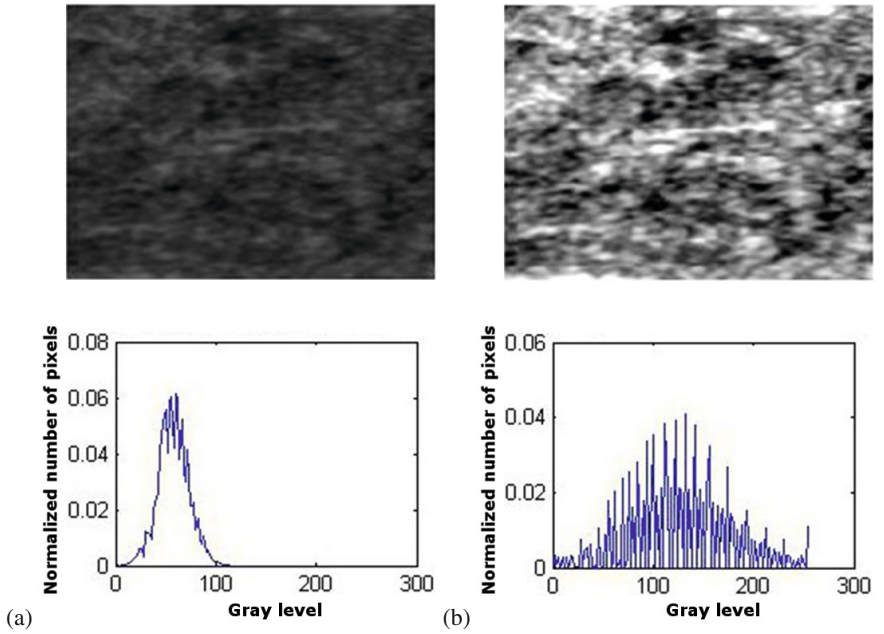


Fig. 3 An example of a sample image before and after normalization: (a) source image and its histogram, (b) normalized image and its histogram

The set of features was obtained on the basis of:

1. Histogram (9 features): histogram mean, histogram variance, histogram skewness, histogram kurtosis, percentiles 1%, 10%, 50%, 90%, and 99%.
2. Gradient (5 features): absolute gradient mean, absolute gradient variance, absolute gradient skewness, absolute gradient kurtosis, percentage of pixels with nonzero gradient.
3. Run length matrix (5 features $\times$ 4 various directions): run length nonuniformity, grey level nonuniformity, long run emphasis, short run emphasis, fraction of image in runs.
4. Coocurrence matrix (11 features $\times$ 4 various directions $\times$ 5 between-pixels distances) angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy.
5. Autoregressive model (5 features): parameters Θ_1, Θ_2, Θ_3, Θ_4, standard deviation.
6. Haar wavelet (24 features): wavelet energy (features are computed at 6 scales within 4 frequency bands LL, LH, HL and HH).

4 Classifier Construction and Testing

4.1 Classifier Based on Decision Tree Induction

Description. An experiment was carried out using a CART algorithm [3]. The 10-fold cross-validation method was used to evaluate the model quality, due to the limited number of cases in the analyzed data set. Due to the aforementioned initial stage of research, procedures for eliminating the irrelevant (with respect to the purpose of classification) descriptive attributes were omitted, and it was considered as appropriate to evaluate their validity in the process of decision tree parameterization. There were applied the equal costs of misclassification, Gini coefficient as a criterion of goodness of fitting and a priori probability of occurrence of two analyzed categories estimated on the basis of their number. The minimum number of cases in the node was equal to 13 and the maximum number of nodes was equal to 1000. The pruning, when the misclassification occurs as a stopping rule, has been chosen [12].

Results. Six decision trees were created as a result of the pruning. The misclassification costs of these trees evaluated by the resubstitution and 10-fold cross-validation methods in Figure 4 have been shown. The graph of the optimal tree (no. 4), due to the size of the classification error (cost) on the test set, was shown in Figure 5, and the classification matrix of cases of the analyzed data set was presented in Table 1. Table 2 shows the importance of descriptive attributes due to the accepted classification purpose, using the optimal tree.

Discussion. The selected optimal model, CART decision tree no. 4, classifies the cases of patients who were diagnosed with Hashimoto's disease with an accuracy of 89.41%, and it assigns incorrectly to this category 5.56% of healthy patients. The false-negative classification error twice higher than the false-positive error is not favorable from a medical point of view (a positive classification is if the model recognizes the disease correctly). However, it should be noted that this disadvantage can be resolved in further research by differentiating the costs of misclassification in the phase of model construction. A classification error below 20% was estimated by using the 10-fold cross-validation method on an independent test set. Taking into account the typical standards of machine learning, this result gives the possibility for reliable application of presented methodology in the planned advisory system that would support the physician in the diagnosis of Hashimoto's disease.

Table 1 Classification matrix for CART

		Predicted		
Parameter	Observed	Sick	Healthy	Total in row
Frequency	Sick	76	9	85
Percentage [%]		89.41	10.59	100
Frequency	Healthy	3	51	54
Percentage [%]		5.56	94.44	100

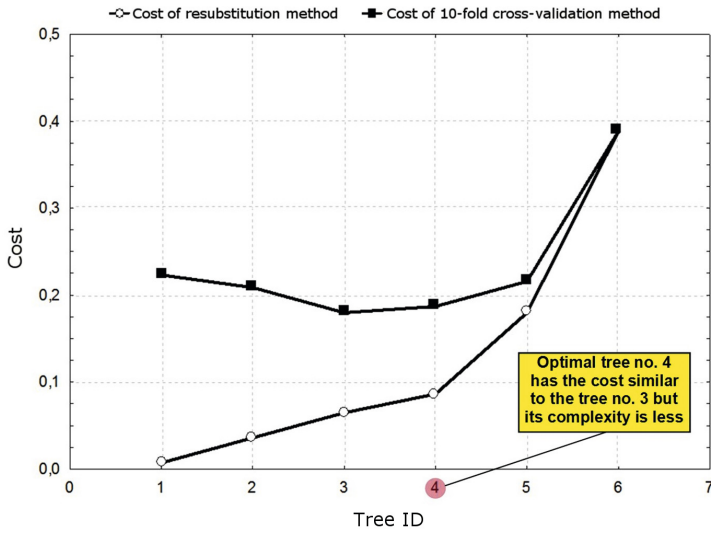


Fig. 4 Sequence of the misclassifying costs

Table 2 Importance of descriptive attributes for CART

No.	Name	Meaning	Importance
1	Theta2	Par. of the autoregression model	100
2	WavEnLH_s-6	Wavelet energy	96
3	WavEnHL_s-6		87
4	WavEnHH_s-5		87
5	WavEnLL_s-4		84
6	S(5,5)DifVarnc	Difference variance	76
7	S(5,-5)DifVarnc		74
8	WavEnLH_s-3	Wavelet energy	73
9	WavEnLL_s-1		73
10	WavEnLL_s-2		71

The analysis of the tree parameters raises also a presumption that will be reviewed at the next stage of research. Namely, node no. 9, which contains a relatively large number of misclassified cases (7 cases of sick were misclassified as healthy), may contain very interesting (from a medical point of view) cases of disease symptoms in its early stage. It is also worth noting the fact that, from the point of view of classification, out of 283 discriminant features, only those features have the greatest importance purpose that were obtained by using the discrete wavelet transform. Seven out of ten best descriptors were obtained by using this method.

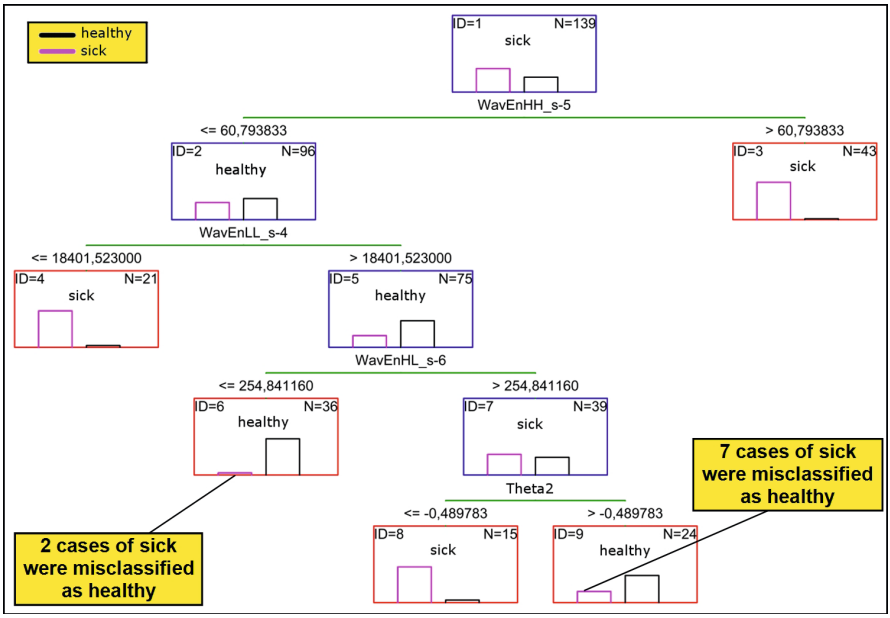


Fig. 5 Optimal decision tree no. 4 (number of divisions equals 4, number of terminal nodes equals 5)

4.2 Classifier Based on MLP Neural Network

Description. The classifier model was constructed by means of the multilayer perceptron neural network. The 10-fold cross-validation method was used to evaluate the model quality, due to the limited number of cases in the analyzed data set. For the same reason, the number of input attributes was limited by using the network sensitivity analysis. The perceptrons were trained by means of the backward error propagation method.

Results. The results of the classification for the best model MLP 6-8-1 (6 input neurons, 8 neurons in the hidden layer and 1 output neuron) are shown in Table 3. This model was selected by using the 10-fold cross-validation method. Table 4 shows the importance of the descriptive attributes which have been selected by means of the network sensitivity analysis.

Discussion. The selected optimal model classifies the cases of patients who were diagnosed with Hashimoto's disease with an accuracy of 89.41%, and assigns, incorrectly to this category, 38.89% of healthy cases. A very low false-negative classification error is very favorable from a medical point of view. The classification error estimated by the 10-fold cross-validation method is at the level of 27.69%. This fact and the classification results indicate that, as a result of a proper correction at the stage of the network architecture selection and its learning process, the artificial

Table 3 Classification matrix for WMA

		Predicted		
Parameter	Observed	Sick	Healthy	Total in row
Frequency	Sick	76	9	85
Percentage [%]		89.41	10.59	100
Frequency	Healthy	21	33	54
Percentage [%]		38.89	61.11	100

Table 4 Importance of descriptive attributes for WMA

No.	Name	Meaning	Importance
1	S(5,5)DifEntrp	Difference entropy	2.10
2	S(0,2)Correlat	Correlation	1.94
3	S(3,0)AngScMom	Angular second moment	1.31
4	WavEnHL_s-2	Wavelet energy	1.06
5	WavEnLH_s-6		1.04
6	WavEnHL_s-1		1.01

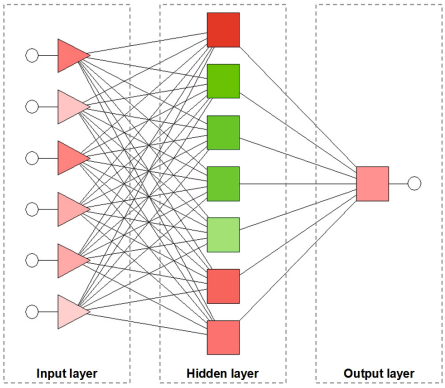


Fig. 6 The best model MLP 6-8-1. The level of signal, stimulating the neurons, varies from −1 to +1 for the input and hidden layer, and it varies from 0 to +1 for the output layer

neural networks can also be a valuable element of the set of classifiers. Such a set can be built, for example, by means of the Weighted Majority Algorithm (WMA).

It is also worth noting the fact that out of 283 discriminant features, only those features are especially useful for the network construction that were obtained by means of the discrete wavelet transformation. Three out of six best descriptors were obtained by this method.

5 Conclusions

The objective of the study was the classification of selected fragments of the thyroid ultrasound images. There were 54 samples of the healthy thyroid, and 85 samples diagnosed with Hashimoto's disease. The images were normalized and then a set of 283 discriminant features was extracted. On this basis, two classifiers have been constructed by using the decision tree induction and the multilayer perceptron neural network. The tests performed on the available image samples have shown that the built models can be successfully used as a part of an advisory system that would support a physician in the diagnosis of Hashimoto's disease. The same classification accuracy of sick cases at the level of 89.41% by using two different methods has been obtained. It should be noted that the number of misclassified cases of sick is identical (9) for both methods and only two cases are the same. A group of sick cases which was misclassified as healthy seems to be very significant for further research. According to the authors, this group of cases may contain information interesting for detection of the early stage of Hashimoto's disease. The aim of further research will be to verify of this thesis and compare the classification results obtained by means of the CART algorithm and the multilayer perceptron neural network with the results achieved by means of other machine learning methods.

References

1. The home page of MaZda, `http://www.eletel.p.lodz.pl/programy/cost/progr_mazda.html`
2. Bastanfard, M., Jalaeian, B., Jafari, S.: Analysis of sonogram images of thyroid gland based on wavelet transform. International Journal of Computer, Information, Systems and Control Engineering 1(3), 194–197 (2007)
3. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Chapman & Hall, Boca Raton (1993)
4. Chang, C.Y., Chen, S.J., Tsai, M.F.: Application of support-vector-machine-based method for feature selection and classification of thyroid nodules in ultrasound images. Pattern Recognition 43(10), 3494–3506 (2010)
5. Chang, C.Y., Huang, H.C., Chen, S.J.: Thyroid nodule segmentation and component analysis in ultrasound images. In: Proceedings of the 2009 APSIPA Annual Summit and Conference, Sapporo, Japan, pp. 910–917 (2009)
6. Chang, W.W., Yeh, W.C., Huang, P.C.: A hybrid immune-estimation distribution of algorithm for mining thyroid gland data. Expert Systems with Applications 37(3), 2066–2071 (2010)
7. Dogantekin, E., Dogantekin, A., Avci, D.: An automatic diagnosis system based on thyroid gland: ADSTG. Expert Systems with Applications 37(9), 6368–6372 (2010)
8. Dogantekin, E., Dogantekin, A., Avci, D.: An expert system based on generalized discriminant analysis and wavelet support vector machine for diagnosis of thyroid diseases. Expert Systems with Applications 38(1), 146–150 (2011)
9. Fauci, A., Braunwald, E., Isselbacher, K., Wilson, J., Martin, J., Kasper, D., Hauser, S., Longo, D. (eds.): Harrison's Principles of Internal Medicine, vol. 3. McGraw-Hill, New York (1998)

10. Haralick, R.: Statistical and structural approaches to texture. Proceedings of the IEEE 67(5), 786–804 (1979)
11. Haralick, R., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics SMC-3(6), 610–621 (1973)
12. Hill, T., Lewicki, P.: STATISTICS: Methods and Applications. StatSoft, Tulsa (2007)
13. Hu, Y., Dennis, T.: Textured image segmentation by context enhanced clustering. IEE Proceedings - Vision, Image and Signal Processing 141(6), 413–421 (1994)
14. Iakovidis, D.K., Savelonas, M.A., Karkanis, S.A., Maroulis, D.E.: A genetically optimized level set approach to segmentation of thyroid ultrasound images. Applied Intelligence 27(3), 193–203 (2007)
15. Katsigiannis, S., Keramidas, E., Maroulis, D.: A contourlet transform feature extraction scheme for ultrasound thyroid texture classification. Engineering Intelligent Systems 18(3), 171–182 (2010)
16. Keles, A., Keles, A.: ESTDD: Expert system for thyroid diseases diagnosis. Expert Systems with Applications 34(1), 242–246 (2008)
17. Keramidas, E.G., Iakovidis, D.K., Maroulis, D., Dimitropoulos, N.: Automatic measurement of thyroid gland. In: Proceedings of the Panhellenic Conference on Informatics (PCI 2007), Patra, Greece, pp. 49–56 (2007)
18. Keramidas, E.G., Iakovidis, D.K., Maroulis, D., Karkanis, S.: Efficient and effective ultrasound image analysis scheme for thyroid nodule detection. In: Kamel, M.S., Campilho, A. (eds.) ICIAR 2007. LNCS, vol. 4633, pp. 1052–1060. Springer, Heidelberg (2007)
19. Keramidas, E.G., Maroulis, D., Iakovidis, D.K.: Tnd: A thyroid nodule detection system for analysis of ultrasound images and videos. Journal of Medical Systems 36(3), 1271–1281 (2012)
20. Kodaz, H., Ozsen, S., Arslan, A., Gunes, S.: Medical application of information gain based artificial immune recognition system (AIRS): Diagnosis of thyroid disease. Expert Systems with Applications 36(2, Part 2), 3086–3092 (2009)
21. Koprowski, R., Wrobel, Z., Zieleznik, W.: Automatic ultrasound image analysis in hashimoto's disease. In: Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A., Kittler, J. (eds.) MCPR 2010. LNCS, vol. 6256, pp. 98–106. Springer, Heidelberg (2010)
22. Lerski, R., Straughan, K., Schad, L., Boyce, D., Blml, S., Zuna, I.: MR image texture analysis - an approach to tissue characterization. Magnetic Resonance Imaging 11(6), 873–887 (1993)
23. Ligeza, A.: Logical Foundations for Rule-Based Systems. AGH Krakow (2005)
24. Quinlan, J.: Induction of decision trees. Machine Learning 1(1), 81–106 (1986)
25. Rumelhart, D., McClelland, J. (eds.): Parallel distributed processing. MIT Press, Cambridge (1986)
26. Šára, R., Smutek, D., Sucharda, P., Svacina, S.: Systematic construction of texture features for hashimoto's lymphocytic thyroiditis recognition from sonographic images. In: Quaglini, S., Barahona, P., Andreassen, S. (eds.) AIME 2001. LNCS (LNAI), vol. 2101, pp. 339–348. Springer, Heidelberg (2001)
27. Seabra, J.C.R., Fred, A.L.N.: Towards the development of a thyroid ultrasound biometric scheme based on tissue echo-morphological features. In: Fred, A., Filipe, J., Gamboa, H. (eds.) BIOSTEC 2009. CCIS, vol. 52, pp. 286–298. Springer, Heidelberg (2010)
28. Supriyanto, E., Arif, N.M., Rusli, A.H., Humaimi, N.: Semi-automatic thyroid area measurement based on ultrasound image. In: Proceedings of the 15th WSEAS International Conference on Computers, Corfu Island, Greece, pp. 228–233 (2011)

29. Švec, M., Šára, R., Smutek, D.: On reproducibility of ultrasound image classification. In: Marques, J.S., Pérez de la Blanca, N., Pina, P. (eds.) IbPRIA 2005. LNCS, vol. 3523, pp. 439–446. Springer, Heidelberg (2005)
30. Temurtas, F.: A comparative study on thyroid disease diagnosis using neural networks. Expert Systems with Applications 36(1), 944–949 (2009)
31. Venables, W., Ripley, B.: Modern Applied Statistics with S-PLUS. Statistics and Computing. Springer, New York (1997)
32. Yeh, W.C.: Novel swarm optimization for mining classification rules on thyroid gland data. Information Sciences 197, 65–76 (2012)
33. Zurada, J.M.: Introduction to Artificial Neural Systems. PWS Publishing Company (1992)

A Fuzzy Logic Approach to Measure of Weight Status and Central Fatness in Adults and Adolescents

Tadeusz Nawarycz, Krzysztof Pytel, Wojciech Drygas, Maciej Gazicki-Lipman, and Lidia Ostrowska-Nawarycz

Abstract. Body mass index (BMI) as a measure of body weight status (BWtS) and waist circumference (WC) and/or waist-to-height ratio (WHtR) as a measure of central fatness (CF) are the most widely used indices of the cardio-metabolic risks (CMR) in prophylactic and clinical practice. The chapter presents, coherent for adults and adolescents, fuzzy models of both BWtS and CF as well as their usefulness in the assessments of CMR. For the construction of a membership function (MF), the Zadeh's Extension Principle (EP) and mapping of the BMI fuzzy sets into adequate CF fuzzy sets using different transformation functions are applied. Taking advantage of the results of a screening study, the CF membership functions for the adult population of the city of Lodz (Poland) as well as for its adolescent population are presented. MF design based on the EP theory is a useful methodology for an estimation of the CF fuzzy subsets and, consequently, for a better assessment of CMR and other medical diagnostic applications.

Tadeusz Nawarycz
Chair of Experimental and Clinical Physiology, Department of Biophysics, Medical University of Lodz, Poland
e-mail: tadeusz.nawarycz@umed.lodz.pl

Krzysztof Pytel
Faculty of Physics and Applied Informatics, University of Lodz, Poland
e-mail: kpytel@uni.lodz.pl

Wojciech Drygas
Department of Social and Preventive Medicine, Medical University of Lodz, Poland

Maciej Gazicki-Lipman
Chair of Experimental and Clinical Physiology, Department of Biophysics, Medical University of Lodz, Poland
e-mail: maciej.gazicki-lipman@umed.lodz.pl

Lidia Ostrowska-Nawarycz
Chair of Experimental and Clinical Physiology, Department of Biophysics, Medical University of Lodz, Poland
e-mail: lidia.ostrowska-nawarycz@umed.lodz.pl

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*, Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_3

1 Introduction

Obesity is an important risk factor of cardiometabolic diseases (CMD), and a primary public health problem in many countries of the world [1]. The trend in obesity is especially alarming among children and adolescents. The annual rate of increase in the prevalence of childhood obesity has been growing steadily, and the current rate is 10 times that observed in the 1970s. This contributes to the obesity epidemic in adult population and creates a growing health challenge for the next generation [2]. During the last century, global eating habits shifted from fiber-rich foods to an obesity prone Western diet. In 2010, obesity was affecting about one third of US adult population aged 20 and older [3]. Pol-MONICA (Multinational Monitoring of Trends and Determinants in Cardiovascular Diseases) study has shown that overweight is exhibited by about 67% of the Polish population, while approximately 30% of women and 20% of men suffer from obesity [4].

Although body mass index (BMI) is routinely used to evaluate body weight status (BWtS) and to measure overall obesity, there is now an evidence showing that indices of central fatness (CF) make better predictors of cardio-metabolic complications [5], [6]. An ectopic fat accumulation occurs in muscle, liver, and heart, as well as "fat infiltration" in other tissues. The end results are cellular dysfunction, apoptosis and consequent pathologic stress on organs, such as heart failure, β-cell dysfunction in the pancreas, peripheral insulin resistance in both liver and muscle tissue[7], [8].

Central fatness evaluation based on exact measurements of visceral fat is very difficult and it is only possible to be carried out with the help of complex and expensive imaging techniques such as computer tomography (CT), magnetic resonance imaging (MRI) or dual-energy X-ray absorptiometry (DEXA) [9], [10]. In practice, an unfavorable distribution of adipose tissue and a presence of CF are usually estimated by indirect methods using simple anthropometric measurement.

Recently, anthropometric measures that reflect CF, such as waist circumference (WC) and waist-to-height ratio (WHtR calculated as a the ratio of WC to body height), have been, both in adults and in children, related to an increased risk of all-cause mortality [11], [12]. Indices of CF are suggested as being superior to BMI in predicting cardiommetabolic risk (CMR) and all-cause mortality [13], [14]. An excess of visceral fat surrounding abdominal organs substantially increases a risk of life-threatening diseases such as type 2 diabetes and heart diseases [6], [15]. These diseases are the most common cause of death and have the greatest impact on the financial burden of both individual and public health system [16].

There are a number of etiological factors producing obesity and these include both genetic, endocrine alterations and environmental factors and hence it is classified as a multifactor disorder [17]. In addition, the distinct and changing economic, social, cultural, and environmental factors play a significant role in the onset of obesity [18]. All causes of obesity are characterized by their imprecision and their uncertainty.

Current recommendations establish sharp boundaries between terms like overweight, obesity, etc. They are often of a contractual nature of a consensus obtained

in the form of recommendation. Issues related to the assessment of CF belong to those unspecific problems of body composition which could be well approached by a theory of fuzzy sets [19], [20], [21].

The chapter presents, coherent for adults and adolescents, fuzzy models of both BWtS and CF as well as their usefulness in the assessments of CMR. For the construction of a membership function (MF), the Zadeh's Extension Principle (EP) and mapping of the BMI fuzzy sets into adequate CF fuzzy sets, using different transformation functions, are applied. Results of the respective studies, curried out on the population of the city of Lodz (Poland), were used as realistic example of the theoretical considerations.

This chapter consists of the following main subsections: Introduction (1), Body weight status and central fatness - current recommendations (2), A fuzzy logic approach to the evaluation of obesity status (3), and Discussion and Conclusions (4).

2 Body Weight Status and Central Fatness – Current Recommendations

2.1 Body Weight Status According to BMI

Both in clinical practice and in epidemiological studies, BWtS is most frequently determined on the basis of the BMI value. BMI is a simple index of "weight-for-height" calculated as weight expressed in kilograms divided by squared height expressed in meters, [kg/m^2]. BMI values are age-independent and identical for both sexes. The classification of BMI values recommended for adults [1], [22] and for children/adolescents [23] is presented in Table 1.

Table 1 Classification of the body weight status according to BMI.

Classification	**Children** [Percentile]	**Adults** [kg/m^2]
Underweight	$<$ P5	$<$ 18.5
Normal (NO)	P5 - P84	18.5 - 24.9
Overweight (OW)	P85 - P94	25 - 29.9
Obese (OB):	$\geq$ P95	$\geq$ 30
Class 1	P95 - P98	30 - 34.9
Class 2	$\geq$ P998	35 - 39.9
Class 3		$\geq$ 40

Adults with normal BWtS are those whose BMI is higher than 18.5 [kg/m^2] and lower then 25 [kg/m^2]. Overweight (OW) is defined as BMI of between 25 and 30 [kg/m^2] and obesity (OB) is defined as BMI of $\geq$ 30 [kg/m^2] [1]. However, BMI is

the one BWtS index and it has several limitations: it does not accurately estimate the body fatness (fat mass and muscle mass are not distinguished), it does not take into account physiological differences between males and females adiposity and it does not always budge in response to lifestyle change. BMI is a measure of relative weight and it is only a proxy indicator of body fatness. It does not provide any information about the body fat distribution or a body shape. Despite the above-mentioned limitations, BMI is a popular measure because it constitutes a relatively easy, inexpensive an non-invasive proxy measure of an excess global body fatness [6], [24].

In the case of children and adolescents, the BMI cut-off points - in a way similar to other developmental parameters - are more difficult to define without a controversy and still are an issue a considerable debate.

2.2 Categorization of Health Risk Using BMI and Waist Circumference

Recently, the WC and WHtR indices have been successfully used as an effective measure of CF both in adult and in children or adolescent populations.

In 1995, Lean et al. [25] proposed that the WC alone could be used as a measure for indicating need for weight management and to define suitable values for risk groups. Based on population study concerning a relationship between BMI and WC, they suggested Action Level 1 (AL1) for a WC $\geq$ 94 cm for men and WC $\geq$ 80 cm for women and Action Level 2 (AL2) for a WC $\geq$ 102 cm for men and WC $\geq$ 88 cm for women [35]. Currently, the WC cut-off points in accordance with the AL1 and AL2 are recommended as the definitions of CF (abdominal obesity) by the International Diabetes Federation (IDF) and by the American experts (NCEP-ATP III : Third Report of the National Program of Cholesterol Education) [26], [27].

The WHO recommended that an individual relative risk could be more accurately classified using both BMI and two stage of AL for WC (Table 2).

Table 2 Combining BMI and WC action levels for male and female (M/F) populations to classify the cardio-metabolic (CMD) disease risk [28], [29].

	Waist Circumference [cm]	
BMI [kg/m^2]	(AL1 - AL2) M/F 94/80 - 102/88	(AL2) M/F $>$ 102/88
$<$ 18.5	-	-
18.5 - 24.9	-	Increased
25 - 29.9	Increased	High
$\geq$ 30	High	Very high

In the case of children and adolescents, the criteria for CF diagnosis still are not clearly defined. Some authors suggests that a combination of BMI and WC should also be used in clinical settings to evaluate the presence of elevated health risk among children and adolescents [30].

Different approaches have been used to evaluate WC cut-off values for children and adolescents, including extrapolations from adult cut-offs, receiver operating characteristic (ROC) analysis or specific arbitrary percentiles. In practice, the individual CF diagnosis in children and adolescents is usually performed based on representative regional (national) reference systems with the WC 90th percentile (P90) as the cutoff point. Jelliffe and Janssen [31] suggested, as the cut-off points for all MetS risk factors, the percentile levels at the age of 18 corresponding to the adult criteria. This approach eliminates the arbitrary choice of cut-off points and ensures their continuity with the stable, legitimate criteria associated with the adult population.

In the present work, we prefer an approach based on AL concept as a measure that is both universal (for different ethnicity) and harmonized (for adults and pediatric subjects), used in the task of an evaluation of obesity related risks.

2.3 Central Fatness Indices – BMI Relationships

On the basis of the AL concept for WC, using the current recommendations, the following linear WC - BMI relations for men (M) and women (F) can be specified:

- f1: WC = f1(BMI); linear equations based on the NCAP- ATP III / IDF criteria [26], [27].

$$f1: \begin{cases} WC_M = 1.6 \cdot BMI + 54 \\ WC_F = 1.6 \cdot BMI + 40 \end{cases} \tag{1}$$

Similarly, in the case of WHtR, by taking the value of 0.5 as AL1 and the value of 0.6 as AL2, we can determine a gender- and age- independent linear relationship in the form:

- f2: WHtR = f2(BMI); linear equation prepared on the basis on Ashwell M., recommendations [32]:

$$f2: WHtR_{M/F} = 0.02 \cdot BMI \tag{2}$$

Graphical representations of equations (1) and (2), together with the selected risk areas, are shown in Figure 1.

It should be noted that, due to the dichotomous character of the classification systems based on BMI and the CF indices as well as the problem of their cut-off points, the CMR inference table (eg. Table1) constitute rather a qualitative description of the cardio-metabolic complications.

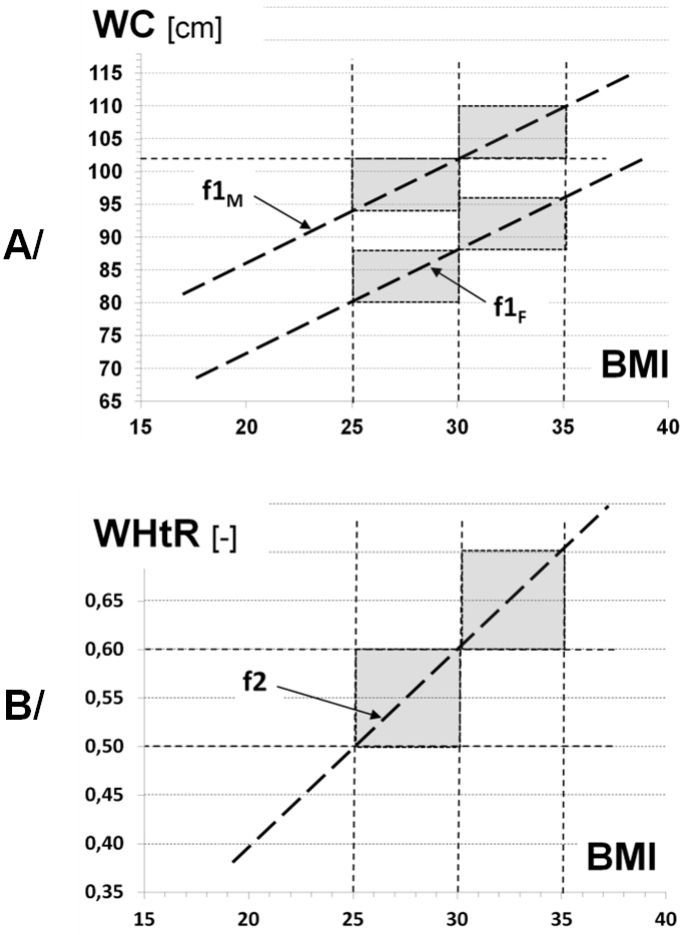


Fig. 1 Linear relation between central fatness indices and BMI. A/ $f1_M$, $f1_F$: BMI - WC (waist circumference) relationship for male (M) and female (F) respectively adopted from IDF and NCEP-ATPIII criteria [26], [27]. B/ f2: BMI - WHtR linear relationship created according to Ashwell M. et al. suggestions [32]

In the fuzzy approach of the obesity related health risk, the linear functions f1 and f2 can be regarded as a mapping functions of the BMI fuzzy sets into adequate CF fuzzy sets.

3 A Fuzzy Logic Approach to the Evaluation of Obesity Status

3.1 The Extension Principle Theory

The extension method proposed by Zadeh, usually called Extension Principle (EP) only, is one of the basic ideas used to process the extension of the classical mathematical concepts into fuzzy ones [33], [34]. Let us consider two crisp sets X and Y and f a mapping from X to Y, $f : X \to Y$. Let A be a fuzzy subset of X, $A \in X$ (Figure 1).

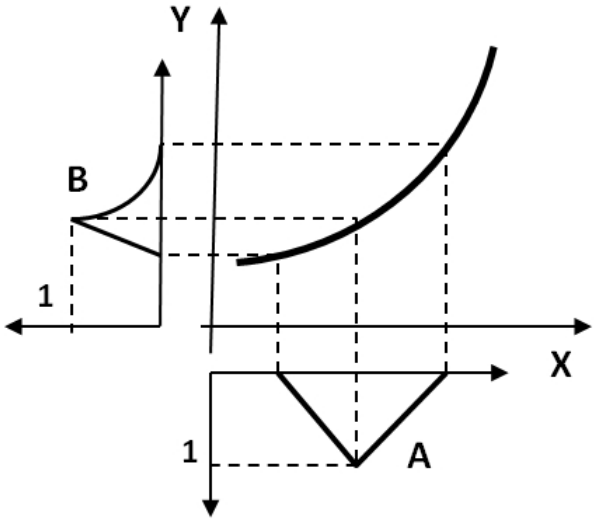


Fig. 2 Illustration of the Extension Principle for a mapping a single variable (explanations in text)

So, the EP allows one to built an image of A under the crisp, thus mapping f as a fuzzy set B:

$$B = f(A) = \{(y, \mu_B(y) \mid y = f(x), x \in X\} \tag{3}$$

whose membership function is given by:

$$\mu_B(y) = \begin{cases} sup\ \mu_A(x),\ if\ f^{-1}(y) \neq \emptyset \\ \\ 0,\ if\ f^{-1}(y) = \emptyset \end{cases} \qquad for\ all\ y \in Y \tag{4}$$

where: $f^{-1}(y)$ denotes the set of all points $x \in X$ such that $f(x) = y$.

In the case of X with an infinite number of elements a fuzzy set B induced by the mapping f can be expressed as:

$$B = f(A) = \int_y \frac{\mu_A(x)}{f(x)} \tag{5}$$

Figure 2 illustrates the EP action considering the mapping $f : X \to Y$ and A a triangular fuzzy set.

3.2 Fuzzy Models of Body Weight Status and Central Fatness

The fuzzy model of the BWtS, based on BMI that has three defined fuzzy subsets (L - low, H - high and VH - very high) with the following MF, has been proposed:

$$L_{BMI} = \begin{cases} 1 & for\ BMI \leq 25 \\ 6 - 0.2BMI & for\ 25 < BMI \leq 30 \\ 0 & for\ BMI > 30 \end{cases} \tag{6}$$

$$H_{BMI} = \begin{cases} 0 & for\ BMI \leq 25 \\ 0.2BMI - 5 & for\ 25 < BMI \leq 30 \\ 7 - 0.2BMI & for\ 30 < BMI \leq 35 \\ 0 & for\ BMI > 35 \end{cases} \tag{7}$$

$$VH_{BMI} = \begin{cases} 0 & for\ BMI \leq 30 \\ 0.2BMI - 6 & for\ 30 < BMI \leq 35 \\ 1 & for\ BMI > 35 \end{cases} \tag{8}$$

A graphic interpretation of MFs of the BMI fuzzy subsets for NO, OW and OB is illustrated in Figure 3.

In the fuzzy approach to the obesity related health risk, the linear functions f1 and f2 can be regarded as mapping functions of the BMI fuzzy sets into adequate CF fuzzy sets.

Example 1: Mapping of the H_{BMI} fuzzy subset based on the proposed model and the fuzzy BMI predefined transformation functions f1 ($H_{BMI} \to H_{WC}$).

Let:

$$H_{BMI} = \{\frac{0}{25} + \frac{0.2}{26} + \frac{0.6}{28} + \frac{1}{30} + \frac{0.6}{32} + \frac{0.2}{34} + \frac{0}{35}\} \tag{9}$$

will be selected fuzzy numbers belonging to H_{BMI} fuzzy set defined by the formula (7). Using the transformation functions f1 (7) defined for men (M) let us calculate appropriate numbers belonging to a fuzzy subset of CF for WC:

$$H_{WC} = \{\frac{0}{f(25)} + \frac{0.2}{f(26)} + \frac{0.6}{f(28)} + \frac{1}{f(30)} + \frac{0.6}{f(32)} + \frac{0.2}{f(34)} + \frac{0}{f(35)}\} \tag{10}$$

and after substituting the values we get:

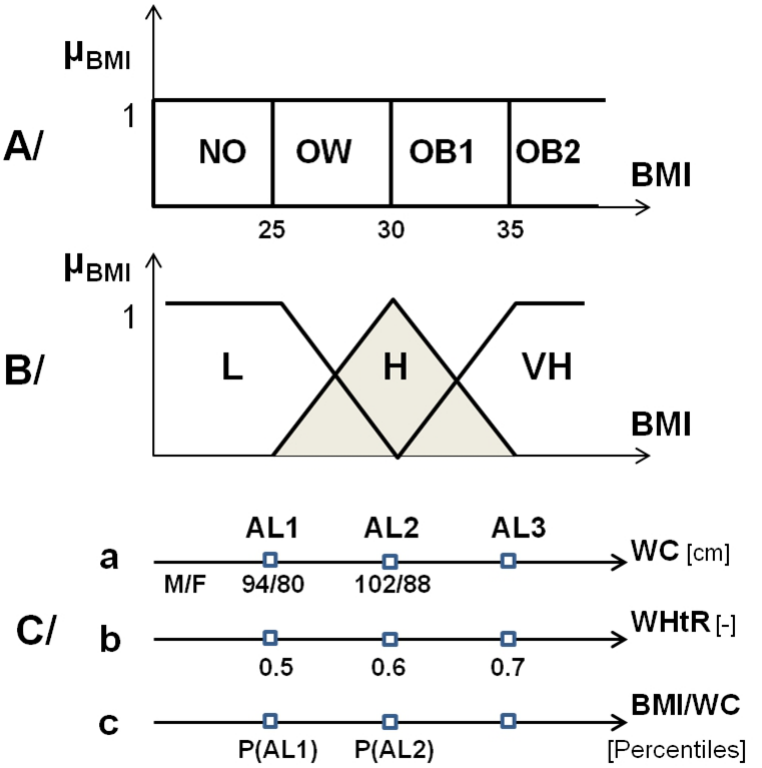


Fig. 3 A graphic interpretation of the MFs for BMI and main action levels (AL) for overall and central fitness indices. A/ Crisp sets of weight status based on actual BMI classification. NO - normal weight, OW - overweight, OB - obesity B/ Fuzzy model of weight status based on BMI action levels. L - low, H - high, VH - very high subsets of the BMI. C/ Action levels of the first (AL1), second (AL2) and third (AL3) stage for WC (a), WHtR (b) and in the case of children and adolescents (c)

$$H_{WC} = \{\frac{0}{94} + \frac{0.2}{95.6} + \frac{0.6}{98.8} + \frac{1}{102} + \frac{0.6}{105.2} + \frac{0.2}{108.4} + \frac{0}{110}\} \tag{11}$$

Similarly, for WHtR and transformation function f2 ($H_{BMI} \rightarrow H_{WHtR}$).

$$H_{WHtR} = \{\frac{0}{0.5} + \frac{0.2}{0.52} + \frac{0.6}{0.56} + \frac{1}{0.6} + \frac{0.6}{0.64} + \frac{0.2}{0.68} + \frac{0}{0.7}\} \tag{12}$$

3.3 Membership Function of the Central Fatness for the Population of Lodz – An Example

In real studies, both the mapping functions and the AL values for CF indices may differ from the functions f1 and f2 shown in the previous subchapter. In the present work, the form of the mapping functions, regional ALs for WC and WHtR as well as the form of CF fuzzy subsets were determined on the basis of the Lodz population study. Matching of the ALs for CF indices can be regarded a process of tuning a system to regional conditions, such as those related to the diversity of socio-economic, ethnic, etc background.

3.3.1 Subjects and Methods

In the present work, the results of epidemiological studies conducted in Lodz in the framework of the project" Early prevention of hypertension, overweight and obesity in children and adolescents - students of Lodz", comprising school age students (14 - 19 years) [35] as well as those of CINDI (Countrywide Integrated Noncommunicable Diseases Intervention) program Lodz, relating to the survey of adults aged 18 - 80 years [36], have been used. The study included a total of 7269 boys and men aged 14 - 80 and 7816 girls and women in the same age range, all residents of the city of Lodz. The three anthropometric parameters: BMI , WC and WHtR were analyzed in all the subjects. The Zadeh's EP and mapping of the BMI fuzzy sets into adequate CF fuzzy sets using adequate regression equations as a transformation functions were applied. In the case of adolescents, the percentile distributions of BMI and WC in terms of gender and age and the LMS method were used [37]. In the case of the 18 - year-old subjects, the specific percentile levels of BMI and WC, corresponding to the typical ALs (cut-off points) for an adult population, were evaluated.

3.3.2 Mapping of the Central Fatness Fuzzy Subsets

On the basis of epidemiological studies of the adult Lodz population, the linear regression equations defined by the WHO-CINDI Programme [36] were designed for WC:

$$f1' : \begin{cases} WC_M = 2.66 \cdot BMI + 22.9 \\ WC_F = 2.27 \cdot BMI + 23.1 \end{cases} \tag{13}$$

In the case of the WHtR the following equations, different for male and female subjects, were obtained:

$$f2' : \begin{cases} WHtR_M = -16 \cdot 10^{-5} \cdot BMI^2 + 25 \cdot 10^{-3} \cdot BMI - 0.012 \\ WHtR_F = -12 \cdot 10^{-5} \cdot BMI^2 + 22.7 \cdot 10^{-3} \cdot BMI + 0.0125 \end{cases} \tag{14}$$

As one sees, the regression relationships between BMI and CF indices for the subjects from the city of Lodz (eq. 9, 10) are different than the corresponding relationships (eq.1, 2). Examples of the regression relationships WC_M = f1' (BMI) (formula 14) and $WHtR_M$ = f2' (BMI) (formula 10) for the men population of Lodz are graphically illustrated in Figure 4.

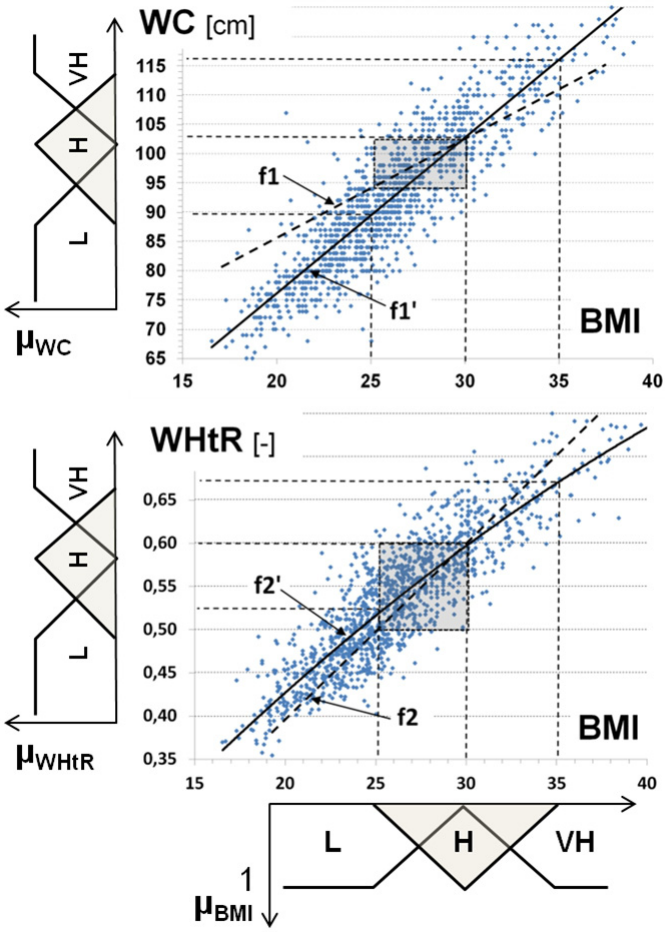


Fig. 4 Examples of mapping the WC (top) and WHtR (bottom) fuzzy sets for the men population of Lodz [18]. The regression equations (f1, f2 - based on recommendations of leading medical organizations; f1', f2' - based on real studies of subjects from the city of Lodz) were used as a transformation functions. The gray rectangles denote the cOW areas based on the NCEP-ATPIII/IDF recommended classification [10], [11]

In the case of pediatric subjects the WC and BMI indices are assessed on the basis of appropriate reference systems (percentile distributions) which are characteristic for the population differentiated by gender, age and ethnicity. The BMI or

Table 3 Waist circumference (WC) and waist circumference-to height ratio (WHtR) cutoff points designated for men (M) and women (F) based on different mapping functions

BMI	WC cut-off [cm]				WHtR cut-off [-]			
[kg/m^2]	(f1)		(f1')		(f2)		(f2')	
	M	F	M	F	M	F	M	F
18.5	83,6	69,6	72,1	65,1			0,396	0,391
25	94	80	89,4	79,8	0,5	0,5	0,513	0,504
30	102	88	102,7	91,2	0,6	0,6	0,594	0,585
35	110	96	116	102,6			0,667	0,659

WC thresholds (ALs or cut-off points) are frequently defined in terms of a specific percentile, on a child/adolescents growth reference.

According to the International Obesity Task Force (IOTF) recommendations [38] confirmed by other authors [31], the "percentile" cut-off points for adolescents 18 years of age should be consistent with the adult criteria. The following are the percentile levels of WC joined with adult criteria (P(AL1) and P(AL2)) for adolescents from the city of Lodz: for boys - P 90.9 and P97.4; for girls - P96.6 and P98.7 respectively. Similarly the percentile levels of BMI joined with adult criteria (BMI=25 [kg/m^2] and BMI=30 [kg/m^2]) are as follows: for boys - P 82.1 and P97.3; for girls - P88.2 and P98.1 respectively. A graph showing the idea of linkage between percentile thresholds and the adult criteria of CF and BMI overweight/obesity definitions is given in Figure 5.

3.3.3 Fuzzy Central Fatness Subsets

WC and WHtR cut - off points designated for male (M) and female (F) populations, based on mapping functions f1, f1', f2, f2', are summarized in Table 3. As we can see, the cut-off points for the fuzzy subsets of CF differ from one another and depend on transformation function. For example, with respect to the triangular MF describing the fuzzy subset H_{WC} of the Lodz inhabitants, both for men and women, the following relationship (Figure 6) is observed:

$$Supp(H_{WC})_{f1} < Supp(H_{WC})_{f1\prime} \tag{15}$$

It is easy to note that this relationship is a consequence of the relationship between the directional coefficients of the regression function: slope (f1') $>$ slope (f1), (equations: (1), (13); Figure 4).

In case of the fuzzy subset H_{WHtR} determined by WHtR, this relation is not so evident (Table 3). Both for men and women from Lodz, the WHtR cut-off points determined with the use of correlation functions f2' were closer to the current expert recommendations described by function f2.

Example 2: Mapping of the H_{BMI} fuzzy subset based on the proposed model into the adequate H_{WC} fuzzy subset ($H_{BMI} \rightarrow H_{WC}$) for men from the city of Lodz and

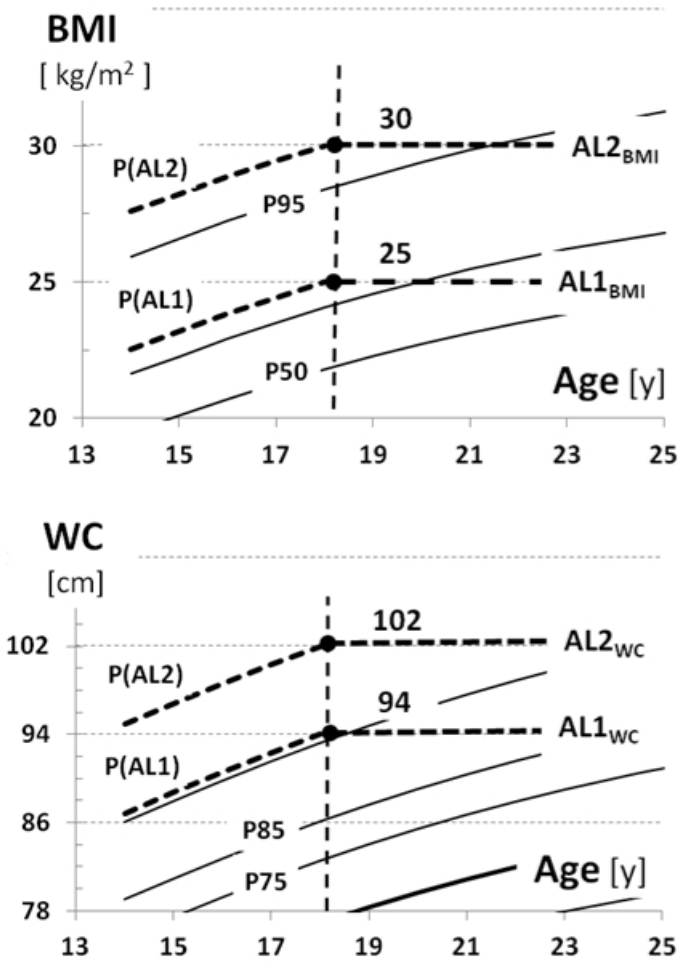


Fig. 5 The BMI (top) and the WC (bottom) percentile curves for boys and young men (14-25y old) from the city of Lodz (Poland). Dotted lines: the thresholds values of BMI (as a measure of BWtS) and WC (as a measure of CF) consistent for youth and adults (explanations in text)

characteristic transformation function f1' (equation 13, Figure 6) gives the following results:

$$H_{WC} = \{\frac{0}{89.5} + \frac{0.2}{92.1} + \frac{0.6}{97.4} + \frac{1}{102.8} + \frac{0.6}{108.1} + \frac{0.2}{113.4} + \frac{0}{116.1}\} \qquad (16)$$

And similarly, for WHtR and transformation function f2' ($H_{BMI} \rightarrow H_{WHtR}$).

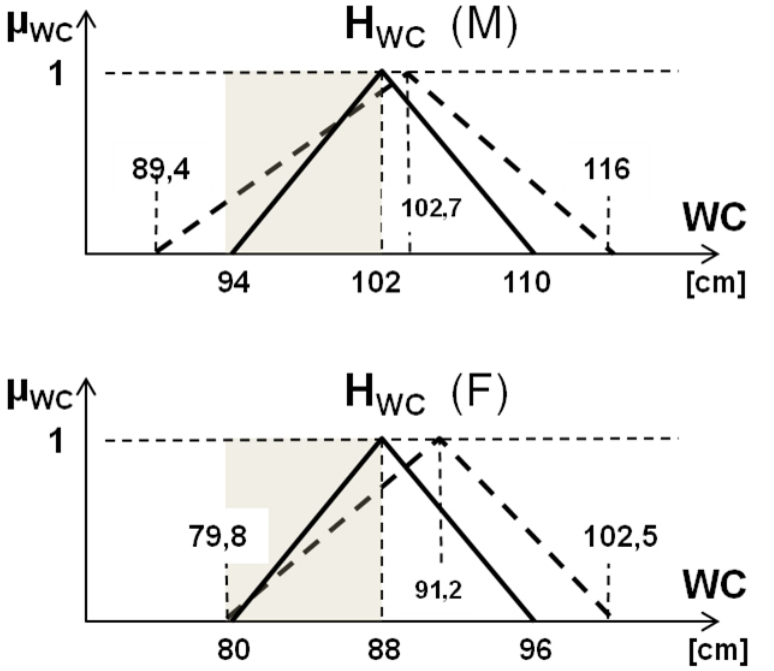


Fig. 6 Graphical form of triangular membership function for the H (high) fuzzy subset of WC for men (upper) and women (bottom). Solid line - developed based on NCEP-ATP III and IDF criteria (function f1); dotted line - developed for the population of Lodz (function f1')

$$H_{WHtR} = \{\frac{0}{0.51} + \frac{0.2}{0.53} + \frac{0.6}{0.56} + \frac{1}{0.59} + \frac{0.2}{0.62} + \frac{0.2}{0.65} + \frac{0}{0.67}\} \qquad (17)$$

The selected fuzzy numbers of the "High" CF fuzzy subsets for WC (equation 16) and WHtR (equation 17) were calculated using real transformation functions (f1 'and f2') specified for the men from the city of Lodz.

4 Discussion and Conclusions

The problem of a reliable assessment of abdominal obesity with the help of simple anthropometric measurements is very important both for public health policy and in individual assessment of CMR. Adipose tissue, especially centrally distributed in the body, meets complex regulatory functions that are difficult to describe unambiguously. The excess of visceral fat is the cause of many cardio-metabolic disturbances, mainly in the form of insulin resistance and secondary hyperinsulinemia [19]. The current CF classifications are based mainly on the arbitrarily defined WC or WHtR cut-off points, which often are not fully corresponding to the characteristic distribution of these indicators within a given population (region).

A proper assessment of the risk of developing obesity- related diseases with the help of fuzzy logic is recently a subject of extensive interest. Fuzzy support systems, in which the question of a proper assessment of the degree of obesity plays a key role, among others, are used in the diagnosis of MetS [39], diabetes [40], bariatric surgery [41] or in an assessment of the health- related fitness [42].

The fuzzy models of BWtS and CF indices, presented in this work, can be made useful in various systems supporting a comprehensive diagnosis and monitoring the risk of complex diseases (syndromes) in which it is essential to note the simultaneous coexistence of several risk factors. Due to the continuous nature of the changes in the risk factors, the use of fuzzy models in a comprehensive risk assessment seems particularly appropriate.

The chapter presents a fuzzy model of CF developed on the basis on such simple anthropometric indices as WC and WHtR. For the construction of a MF, the Zadeh's Extension Principle and mapping of the BMI fuzzy sets into adequate CF fuzzy sets using different transformation functions have been applied. The results obtained for the population of Lodz indicate the potential usefulness of the fuzzy set theory in the evaluation of both, the CF and the CMR. It appears reasonable to apply the Zadeh's Extension Principle for the construction of the MF fuzzy subsets describing the CF, and to use the population-specific correlation function as a mapping function.

The results of our study may be regarded a starting point for the construction of more complex fuzzy inference systems aimed at the assessment of global cardiovascular or cardiometabolic risk, where CF is one of the most important risk factors. Taking into account the growing scale of the adverse health consequences of the obesity epidemic in the world, the artificial intelligence methods, including fuzzy inference systems in particular, should become more and more widely used.

References

[1] WHO, Obesity: preventing and managing the global epidemic. WHO Technical Report Series No.894. World Health Organization, Geneva (2000)

[2] Branca, F., Nikogosian, H., Lobstein, T.: The challenge of obesity in the WHO European region and the strategies for response: Summary. WHO & WHO European Ministerial Conference on Counteracting Obesity. Copenhagen: World Health Organization, Regional Office for Europe (2007)

[3] Flegal, K.M., Carroll, M.D., Kit, B.K., Ogden, C.L.: Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. Journal of the American Medical Association 307(5), 491–497 (2012)

[4] Rywik, S., Broda, G., Piotrowski, W., et al.: Epidemiology of Cardiovascular Disease. Pol-MONICA Program. Kard. Pol (supl. 2), pp. 7–35 (1996) (in Polish)

[5] Larsson, B., et al.: Abdominal adipose tissue distribution, obesity, and risk of cardiovascular disease and death: 13 year follow-up of participants in the study of men born in 1913. Br. Med. J. 288, 1401–1404 (1984)

[6] Wang, Y., et al.: Comparison of abdominal adiposity and overall obesity in predicting risk of type 2 diabetes among men. Am. J. Clin. Nutr. 81, 555–563 (2005)
[7] Hill, M.J., Metcalfe, D., McTernan, P.G.: Obesity and diabetes: lipids, 'nowhere to run to'. Clin Sci (Lond) 116, 113–123 (2009)
[8] Capeau, J.: Insulin resistance and steatosis in humans. Diabetes Metab 34, 649–657 (2008)
[9] Despres, J.P., Ross, R., Lemieux, S.: Imaging techniques applied to the measurement of human body composition. In: Roche, A.F., Heymsfield, S.B., Lohman, T.G. (eds.) Human Body Composition, pp. 149–166. Human Kinetics, Chicago (1996)
[10] Genton, L., Hans, D., Kyle, U.G., Pichard, C.: Dual-energy x-ray absorptiometry and body composition: differences between devices and comparison with reference methods. Nutrition 18, 66–70 (2002)
[11] Savva, S.C., Tornaritis, M., Savva, M.E., Kourides, Y., Panagi, A., Silikiotou, N., et al.: Waist circumference and waist-to-height ratio are better predictors of cardiovascular disease risk factors in children than body mass index. Int. J. Obes. Relat. Metab. Disord. 24, 1453–1458 (2000)
[12] Seidell, J.C.: Waist circumference and waist/hip ratio in relation to all-cause mortality, cancer and sleep apnea. Eur. J. Clin. Nutr. 64, 35–41 (2010)
[13] Lee, C.M., Huxley, R.R., Wildman, R.P., Woodward, M.: Indices of abdominal obesity are better discriminators of cardiovascular risk factors than BMI: A meta-analysis. J. Clin. Epidemiol. 61, 646–653 (2008)
[14] Despres, J.P., Lemieux, I., Bergeron, J., Pibarot, P., Mathieu, P., Larose, E., et al.: Abdominal obesity and the metabolic syndrome: Contribution to global cardiometabolic ris. Arterioscler Thromb Vasc. Biol. 28, 1039–1049 (2008)
[15] Yusuf, S., et al.: Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 364, 937–952 (2004)
[16] Hammond, R.A., Levine, R.: The economic impact of obesity in the United States, Diabetes. Metabolic Syndrome and Obesity: Targets and Therapy 3, 285–295 (2010)
[17] Hillman, G.R.: Fuzzy Logic and Biomedicine, UTMB, The University of Texas Medical Branch (2005)
[18] Monteiro, C.A., Conde, W.L., Popkin, B.M.: Obesity in developing countries: biological and ecological factors-independent effects of income and education on the risk of obesity in the Brazilian adult population. J. Nutr. 131, 881–886 (2001)
[19] Bouharati, S., Bounechada, M., Djoudi, A., et al.: Prevention of Obesity using Artificial Intelligence Techniques. Int. J. 1(9), 146–150 (2012)
[20] Rutkowska, D., Starczewski, A.: Fuzzy Inference Neural Network and Their Applications to Medical Diagnosis. In: Szczepaniak, P., Lisboa, P., Kacprzyk, J. (eds.) Fuzzy System in Medicine, Physica - Verlag, Heidelberg (2000)
[21] Massad, E., Ortega, N.R., Barros, L.C., Struchiner, C.J.: Fuzzy Logic in Action: Applications in Epidemiology and Beyond. Studies in Fuzziness and Soft Computing. Springer (2009)
[22] World Health Organization: Obesity: preventing and managing the global epidemic, WHO Technical Report Series 894. WHO Geneva (2000)
[23] Kuczmarski, R.J., Ogden, C.L., Guo, S.S., Grummer-Strawn, L.M., Flegal, K.M., Mei, Z., et al.: 2000 CDC growth charts for the United States: methods and development. Vital Health Stat. 11, 246, 1-190 (2002)

[24] World Health Organisation. Waist Circumference and Waist-Hip Ratio: Report of a WHO Expert Consultation 2008. World Health Organisation: Geneva (2011)
[25] Lean, M.E., Han, T.S., Morrison, C.E.: Waist circumference as a measure for indicating need for weight management. BMJ 311, 158–161 (1995)
[26] Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP). Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). JAMA 2001, vol. 285, pp. 2486–2497 (2001)
[27] Alberti, K.G., Zimmet, P., Show, J.: IDF Epidemiology Task Force Consensus Group. The metabolic syndrome: A new worldwaide definition. Lancet 366, 1059–1062 (2005)
[28] National Health and Medical Research Council, Clinical practice guidelines for the management of overweight and obesity in adults, adolescents and children in Australia. Melbourne: National Health and Medical Research Council (2013)
[29] Report of a WHO Consultation. Obesity: preventing and managing the global epidemic. WHO Technical Report Series 894, i-253 (2000)
[40] Zhu, S., Heshka, S., Wang, Z., Shen, W., Allison, D.B., Ross, R., Heymsfield, S.B.: Combination of BMI and Waist Circumference for Identifying Cardiovascular Risk Factors in Whites. Obes Res. 12(4), 633–645 (2004)
[31] Jolliffe, C.J., Janssen, I.: Development of age-specific adolescent metabolic syndrome, criteria that are linked to the Adult Treatment Panel III and International Diabetes Federation criteria. J. Am. Coll. Cardiol. 49, 891–898 (2007)
[32] Ashwell, M., Gunn, P., Gibson, S.: Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. Obes. Rev. 13, 275–286 (2012)
[33] Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning I. Information Science 8, 199–251 (1975)
[34] Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning II. Information Science 8, 301–357 (1975)
[35] Ostrowska-Nawarycz, L., Nawarycz, T.: Prevalence of excessive body weight and high blood pressure in children and adolescents in the city of Lodz. Kardiol Pol. 65, 1079–1087 (2007)
[36] Leparski, E., Nussel, E.: Protocol and guidelines for monitoring and evaluation procedure CINDI - Countrywide Integrated Noncommunicable Diseases Intervention Programme, vol. 60. Springer, Heidelberg (1987)
[37] Cole, T.: The LMS method for constructing normalized growth standards. Eur. J. Clin. Nutr. 44, 45–60 (1990)
[38] Cole, T.J., Bellizzi, M.C., Flegal, K.M., Dietz, W.H.: Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ 320, 1240–1243 (2000)
[39] Nawarycz, T., Pytel, K., Rutkowska, D., Ostrowska-Nawarycz, L.: Diagnosis of The Metabolic Syndrome in Children - A Fuzzy Clustering Approach. In: 5th Polish and International PD forum-Conference on Computer Science (PD FCCS 2011), May 16-18 (2011)

[40] Zhu, K.Y., Liu, W.D., Xiao, Y.: Application of Fuzzy Logic Control for Regulation of Glucose Level of Diabetic Patient. Machine Learning in Healthcare Informatics, Intelligent Systems Reference Library 56, 47–64 (2014)

[41] Azevedo, J.L.M., Miyahira, S.A., Leal, L.P., et al.: Fuzzy logic as a decision-making support system for the indication of bariatric surgery based on an index (OBESINDEX) generated by the association between body fat and body mass index. In: Available from Nature Proceedings (2011),
http://hdl.handle.net/10101/npre.2011.5516.1

[42] Nawarycz, T., Pytel, K., Ostrowska-Nawarycz, L.: Evaluation of Health-Related Fitness Using Fuzzy Inference Elements. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2012, Part I. LNCS, vol. 7267, pp. 301–309. Springer, Heidelberg (2012)

Reconstruction of Human Motion Trajectories to Support Human Gait Analysis in Free Moving Subjects

Jaroslav Majerník

Abstract. Understanding of human motion is based on analysis of complex motion patterns and requires systematic study of body mechanics. The technologies used to support analysis of human locomotion has advanced dramatically over past decades and were efficiently applied also in many clinical research studies to help clinicians recognize their patients' motion related health problems. Computerized motion capture systems are used to asses human kinematics in quantitative way. Such systems usually track motion of small either passive or active markers attached to the patients' bodies and offer highly accurate measurements. However, they can not by used out of the laboratory because of sophisticated equipment or they are to expensive to be widely used. Therefore, there are efforts to derive kinematics directly from video records where no special apparel is needed. In our work, the digital image processing techniques were used to develop algorithms for automatic detection of human motion trajectories. Based on these algorithms, the marker–free analysis system has been created and tested in analysis of human gait.

1 Image Processing

Current commercial motion analysis systems use various techniques and algorithms to detect and reconstruct motion trajectories [1, 2]. However, these systems have one common objective that is to derive kinematical parameters precisely characterizing motion of analysed subject. An optimization process, leading to their acquisition arise from the scope of the study and from details that has to be analysed [3]. From the technical point of view, the development of such systems is oriented on increasing of accuracy of motion detection in space while minimising technical demands; increasing of motion trajectories reconstruction rate and reduction of time requirements to process and evaluate relevant data [4]. Already existing systems support phases ranging from subject's selection and preparation, through data capturing and

Jaroslav Majerník
Faculty of Medicine, Pavol Jozef Šafárik University in Košice, Košice, Slovakia
e-mail: jaroslav.majernik@upjs.sk

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*,
Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_4

processing to analysis reports [5]. Their quality depends also on knowledge, experiences and financial resources the team developing such system has.

Despite of recent possibilities and advances obtained in the area of motion video–analysis, there is still no unique methodology nor general algorithm. In general, it is possible to specify several approaches to track motion in picture. These include:

- *Detection of motion.* The simplest detection registering each identifiable motion. It can be realized using single static video camera.
- *Detection and localization of moving objects.* More complicated detection. Here, the video camera can be static and objects realize their movements, or the video camera is in motion and the objects are static. The complexity arise from the need to determine space localization of objects to be tracked.
- *Detection of moving objects, trajectories of their motion and eventually prediction of their future trajectories.* Computational time–consuming methods that use object adjusting techniques as direct adaptation; adaptation of object's signs; adaptation of specific points representing object (contours) or graph presentation of moving objects with subsequent adaptation of these graphs. The most sophisticated system can process motion information while moving both the video camera and the objects.

In conjunction with reconstruction of three–dimensional (3D) object or its movement in 3D space it is necessary to solve the problem how to derive its properties from two–dimensional (2D) projections acquired in different times [6, 7]. 2D representation of 3D movement defines an area of motion, where each point is specified by velocity vector corresponding to direction of motion, velocity of motion and distance from video camera (observer) [8].

Nowadays, there exist several projects oriented on development of complex "marker–free" motion analysis [9, 10, 11]. These may utilize similarity analysis of static images; motion analysis at higher level; searching for conformity between pairs of two points in sequence of frames etc. Applications of adaptation are also very frequently used approaches.

Whatever algorithm is used in marker–free analysis, the computer vision is used as a tool to work with visual information. This science stimulates human vision using computerized equipment [12, 13]. Processing and identification of motion in real scenes is usually divided into several phases. This division is not uniquely determined and may vary across various sources. In general, realization of individual steps and the level of their realization depends on real application. Process of image processing using computer vision is shown in Fig. 1.

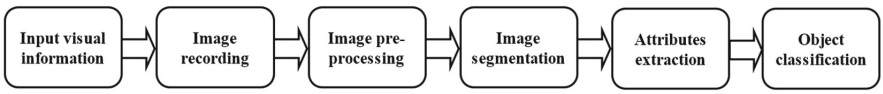


Fig. 1 Principle scheme of image processing based on computer vision

The first block is represented by input scene, i.e. object of interest which has to be transformed into two dimensional function. During recording of the image, the input optical parameters are transformed into electrical signal continuous in time. Here, the brightness or the intensity of thermal or another radiation can be used as input optical parameter. The image can be captured in one or in more spectral bands (for example visible or infrared), whereby it is usually influenced by surrounding illumination and by various physical barriers. The input image is digitized in this phase using matrix representation of the frames and stored on memory to be used in next processing. In the next phase of image pre–processing, the noise as well as irrelevant information can be removed from the image. Here, the data compression is realized to be able to work only with the data needed to process image in the next phases. Various brightness corrections, filtration techniques, zooming and focusing operations are applied to achieve it. Then the image is transformed into uniform areas in the phase of image segmentation. The objects can be separated from background and the process moves to the so called higher level of image processing. The segmentation allows finding the objects that are points of interest regarding next data processing. The block of attributes extraction works only with relevant information while the self shapes of the objects are recognized. These can be described in quantitative way using numerical characteristics or in qualitative way using relations between individual objects. The last phase if object's classification is used to categorize shapes obtained by attributes extraction into specific categories or classes (for example segments of human body).

Despite of highly sophisticated algorithms and latest technologies used in individual phases of image processing chain the understanding of the image and automation of decision process based on information included in image comparing abilities of human brain is still limited and possible only in specific cases.

2 Tracking of Motion in Images

To develop methodology that will be useful for marker–free analysis we respected requirements to minimize disadvantages of recent systems; to minimize difficulties in data processing and to minimize necessary financial expenses. Our attention was oriented mainly on clinical usability as this is our area of interest. Therefore, the final version of the system has to be user friendly with no special technical demands on operating personnel, especially in the phase of image pre–processing and preparation of image signal.

Marker–free systems are developed in several world laboratories and are integrated in various research projects. Because of our requirements we excluded any methods where the additional operations will be needed to modify input images, including background elimination, usage of contrast materials, wearing of special clothing etc. In this way the system will be easy to use and operable in both internal and external conditions. There are some approaches meeting above mentioned criteria and that can be elaborated and implemented in such complex system. These

include for example differential method of motion analysis, motion detection from optical flow or utilization of statistical properties of images.

2.1 Differential Method of Motion Detection

Differential method is used in grey images and its subject manner is to determine the area of motion from non-zero values of differences between levels of grey in two consecutive frames. Formula to determine elements of differential image between consecutive frames F_1 and F_2 is

$$\begin{aligned} d(i,j) &= 0, \quad if \quad |f_1(i,j) - f_2(i,j)| \leq \varepsilon \\ d(i,j) &= 1, \quad if \quad |f_1(i,j) - f_2(i,j)| > \varepsilon \end{aligned} \tag{1}$$

where $f_1(i,j)$ and $f_2(i,j)$ are picture elements in the frames F_1 and F_2, $d(i,j)$ are picture elements of differential image and ε is small positive number determining differential threshold of brightness values. The elements of differential image $d(i,j)$ get the value equal to one if:

1. $f_1(i,j)$ is picture element in moving object and $f_2(i,j)$ is picture element in static background or vice–versa,
2. $f_1(i,j)$ is picture element in moving object and $f_2(i,j)$ is picture element in another moving object or
3. $f_1(i,j)$ is picture element in moving object and $f_2(i,j)$ is picture element in the same part in this moving object.

The results of this method strongly depends on contrast between background and object to be tracked. The disadvantage is that reconstructed trajectories may not be able to reveal the direction of motion. This problem can be solved using so called cumulative differential image, which is created from the sequence of n frames where the referential frame is the first frame F_1. The importance of particular frame can be emphasized by significance coefficients. Then the cumulative differential image will have picture elements $d_{sum}(i,j)$ obtained using equation

$$d_{sum}(i,j) = \sum_{k=1}^{n} a_k |f_1(i,j) - f_k(i,j)|, \tag{2}$$

where a_k are significance coefficients. The current frames may be of higher significance than previous ones to refer importance of current motion. Here, the problem to specify static reference scene may occur, for example when the motion never ends. To determine static reference scene in such cases, it is possible to use self teaching algorithms. Performance of the method can be also increased comparing motion characteristics; using edge detection algorithms or using detection of slowly moving edges. Differential image shows information about the motion, but the characteristics of motion derived from this image are not very plausible.

2.2 Detection of Motion from Optical Flow

Optical flow expresses image changes caused by motion during time period dt. Then, the area of optical flow is given as velocity area, which represents 3D motion of object's points in 2D image.

Let us assume, that there is an continuous dynamic image specified as function of position and time $f(x,y,t)$. This image can be expressed by Taylor's polynomial function

$$f(x+dx,y+dy,t+dt) = f(x,y,t) + f_x dx + f_y dy + f_t dt + O\left(\delta^2\right), \tag{3}$$

where f_x, f_y and f_t are partial derivations of the function f. If the closest adjacent point (x,y) will change its position during time period dt in certain small distance (dx,dy) then it is possible to find dx, dy and dt where

$$f(x+dx,y+dy,t+dt) = f(x,y,t), \tag{4}$$

i.e. if d_x, d_y and d_t are very small, then the high–order relations in the equation (3) are negligible and it results in

$$-f_t = f_x \frac{dx}{dt} + f_y \frac{dy}{dt}. \tag{5}$$

Next, the final objective is to calculate velocities that can be specified by the formula

$$c = \left(\frac{dx}{dt}, \frac{dy}{dt}\right) = (u,v). \tag{6}$$

f_x, f_y and f_t can calculated from $f(x,y,t)$. Then, the motion velocity will be given by

$$-f_t = f_x u + f_y v = grad(f) \cdot c. \tag{7}$$

Grey level difference f_t in the same image position in time t and dt is the result of grey level difference and velocity in this position.

The given formula do not specify velocity vector, but gives component in the direction of brightness gradient. The methd is then reduced to minimize value of quadratic error

$$E^2(x,y) = (f_x u + f_y v + f_t)^2 + \lambda\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right), \tag{8}$$

where the first part is the solution of equation (7); the second one is the smoothness criterion and λ is Lagrange multiplier.

The formula is then reduced to solve differential equations

$$\begin{aligned} \left(\lambda^2 + f_x^2\right) u + f_x f_y v &= \lambda^2 \bar{u} - f_x f_t, \\ f_x f_y u + \left(\lambda^2 + f_y^2\right) v &= \lambda^2 \bar{v} - f_y f_t, \end{aligned} \tag{9}$$

where $\bar{u}$ and $\bar{v}$ are mean values of velocity in direction x and y. Solution of these equations is

$$u = \bar{u} - f_x \frac{P}{D},$$
$$v = \bar{v} - f_y \frac{P}{D}, \tag{10}$$

wher $P = f_x\bar{u} + f_y\bar{v}$ and $D = \lambda^2 + f_x^2 + f_y^2$. Then, the specification of optical flow is based on Gauss–Seidel iteration method using pairs of consecutive dynamic images. Iterations can be slow and in some cases thousands iterations are needed to obtain convergence.

To achieve satisfactory results the optical flow should not be sensitive to lighting changes and to motion of unessential parts (for example shadows). Many problems occur when the static part is lighted by the source of motion and the optical flow is detected. On the other side, soft area rotating in constant lightning offers no optical flow. Errors are caused also by variation of brightness and assumptions of smooth velocity. Except of information about limitations, all calculating errors are distributed across whole system and even small problematic areas are presented this may result in wrong specification of optical flow. Furthermore, the optical flow analysis do not offer direct trajectories of motion. These has to be calculated but the changes between two consecutive frames have to be very small considering time, direction as well as velocity. Therefore, we preferred to not use this method to solve marker–free analysis.

2.3 Statistical Properties of Images

The basis of our method to obtain input parameters for motion tracking algorithm depends on statistical methods. These are applied in almost any area dealing with information. Their sense rise together with improvements of information processing techniques in digital communication, where the emphasis is placed on acquisition, storing and presentation of information in their original or transformed forms. Here, the most of the efforts are oriented on utilization of the most effective ways of data transmission. To make data transfer more efficient, it is possible to improve properties of transmission lines or to realize data compression. Data compression uses elimination of irrelevant and redundant components of video signal. Elimination of irrelevant component of video signal is irreversible process and after encoding it is not possible to restore it. The main criterion to eliminate irrelevant component is visual sensation. On the other hand, the redundant component represents a part of video signal that is preferably eliminated as this part can be restored in the decoding process. The redundancy of visual information is hidden in image statistical characteristics, including for example correlation of picture elements groups, statistical distribution of picture elements etc. These image properties were applied in our motion tracking algorithms in video sequence capturing patient's movements.

Here proposed motion tracking method in video sequence with no passive or active markers attached to the patients' bodies tries to minimize disadvantages of currently available marker–free systems, to bring new of motion analysis and to offer clinicians cheaper, but not worse alternative preserving all advantages of motion analysis. The principle scheme of the system is shown on Fig. 2.

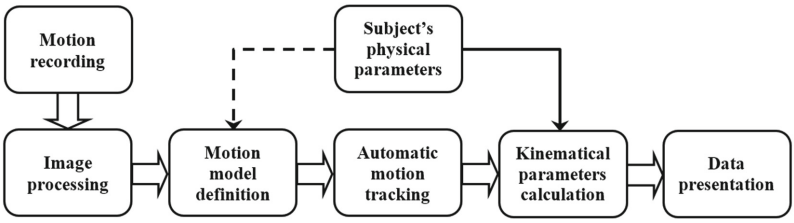


Fig. 2 Principle scheme of marker–free human motion analysis

Motion of the subject is captured using video camera and it is transformed into the computer directly during test or later after one or more motion activity tests are finished. This can be applied also in cases when the tests are realized outside of the laboratory. Video signal is then transformed on sequence of consecutive frames and operator specifies points of interest according to the model to be tracked in one of them. Motion tracking algorithm will search for trajectories across the sequence. Then, these trajectories are used together with anthropometric measures to derive kinematical parameters. Finally, qualitative outputs are presented using numerical and/or graphical form.

3 Recording and Digitizing Image Signal

3.1 Visual Information and Image Discretization

Main motion picture continuous in time, recorded using video camera can be shown in image or in transformed space, where it can be described by deterministic and statistic characteristics. This dynamic continuous image has a character of 3D field and it can be specified in continuous brightness function as

$$l(x,y,t), \tag{11}$$

where x and y are space coordinates and t is the time. An example of continuous image field delimited in plane xy by rectangle with sides labelled as H and V and running in time period T is shown on Fig. 3. Delimitation of space coordinate x is selected with respect to be able to capture required subject's motion in horizontal direction. For example, at least one complete gait cycle should be captured in the case of human gait analysis in sagittal plane. Similarly, vertical dimension V should capture hight of whole subject or at least the complete hight of analysed moving body segments. In such case, the analysed segments will be more visible, but the

relation to whole human body and associated body motions will be missing. The time interval T is usually delimited to the time period of the movement test or to the period when the analysed subject is in the scene.

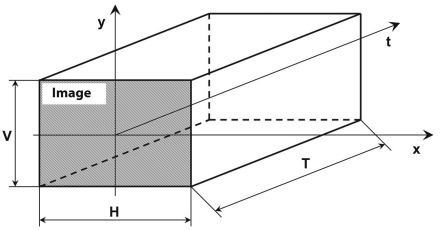


Fig. 3 Continuous image field delimited in space and time coordinates

The time mean image brightness value in the point (x,y) of this continuous image is given as

$$\varepsilon(x,y) = \lim_{T\to\infty}\left[\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} l(x,y,t)\,dt\right], \tag{12}$$

and space mean image brightness value in the time t is

$$\eta(t) = \lim_{H,V\to\infty}\left[\frac{1}{HV}\int_{-\frac{H}{2}}^{\frac{H}{2}}\int_{-\frac{V}{2}}^{\frac{V}{2}} l(x,y,t)\,dxdy\right]. \tag{13}$$

Amount of information in continuous flow is infinite as the number of frames is also infinite. However, perception of human eye is definite, that is why the reduction of redundant information is desirable. It is desired also by digital image processing techniques and realized using limitation of time and spatial discrimination ability.

Time discrimination ability can be reduced by limitation of motion record to finite number of consecutive frames. This strongly depends on the velocity of analysed motion. In slow motions (patients with motion affected diseases) the smaller number of frames will be sufficient (for example 50 Hz) as the motion between two consecutive frames will be small, while in fast motions (athletes) the higher number of frames per second will be required (for example 120 Hz or more) to preserve requirement of small information change between two consecutive frames.

Spatial discrimination ability can be obtained by reduction of the space to finite number of points creating individual frames of motion video record. Discrete image is created as decomposition in image sensors or by discretization in digital systems processing continuous images and it is described using discrete functions. Dynamic discrete image is given by discrete brightness function

$$l'(x,y,t) = \varphi(x,y,t)\,l(x,y,t), \tag{14}$$

where $\varphi(x,y,t)$ is discretization function of continuous image.

Time and spatial discretization of continuous image field using discretization function is shown on Fig. 4.

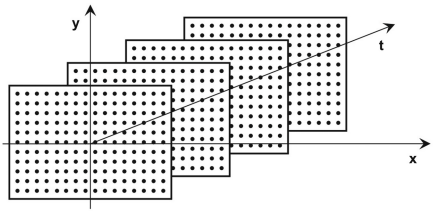


Fig. 4 Time and spatial discretization of continuous image field

3.2 Picture Elements

Digital image of image space is defined by the matrix of $L1 \times L2$ picture elements

$$\underline{x} = [x_{ij}] \begin{matrix} i = 0,1,2,\ldots,L_1 - 1 \\ j = 0,1,2,\ldots,L_2 - 1 \end{matrix} \tag{15}$$

where picture elements x_{ij} are functional values of continuous brightness function in points where the discretization function in given time has non zero values. Process of discretization is realized according to the Shannon's theory, where the smallest detail in digital image has to be at least twice bigger than discretization interval. Selection of optimal image resolution is the most important part of discretization. Low resolution will miss information about details and high resolution will increase further computational demands.

Distance of two picture elements, given for example by coordinates (i, j) and (x, y) should be specified to be able to derive distances, angles, velocities and accelerations. This distance is known as Euclidean distance and can be specified by

$$D_E = \sqrt{(x-i)^2 + (y-j)^2}. \tag{16}$$

In the case of discretized image space, the point vicinity is defined as first and then the distance between picture elements respecting given vicinity. Using square sampling grid it is possible to specify 4–adjacency or 8–adjacency. In the case of 4–adjacency the adjacent picture elements are placed in vertical or horizontal direction. In the case of 8–adjacency the actual picture element has eight possible adjacent picture elements. Then the distances of picture elements are defined as

$$\begin{aligned} D_4 &= |x-i| + |y-j|, \\ D_8 &= max\{|x-i|, |y-j|\}, \end{aligned} \tag{17}$$

where D_4 is the distance for 4–adjacency and D_8 for 8–adjacency.

3.3 Colour Composition of Images

Colour or image colourfulness is significant carrier of information and operations with colour components of picture elements are also included in designed motion tracking algorithm. Various colour models were created to determine exact colour shade. These models distinguished according used colour components. Our system includes several colour models to allow comparison of motion tracking results. The most frequently used models include

- RGB – additive model used in majority of display units (monitors, TVs) where one visible picture element consists of three colour components (red, green and blue). Its combination creates optically different colours. If RGB values are equal to 0,0,0 the resulting colour is back. Increasing values change the output colour until maximal values are reached (1,1,1) that gives white colour. If the values of individual components are equal then there are grey shades on the output.
- CMY – subtractive colour system used primarily in printers and photographs. Here the individual colour (cyan, magenta and yellow) are subtracted from initial white. Printers use extension of the system – CMYK, where K represents black and is used to print black texts.
- HSI – the model in which the individual components are not presented by colours, but by three properties – hue, saturation and intensity.
- YUV – model applied in TV techniques rather than in image processing. Y is the luminance (brightness) component while U and V are the chrominance (colour) components.
- Grey scale – some application process grey images, i.e. colours ranging from white through grey shades to black. Usually 256 shades are defined using 8 bits' colour depth.

4 Image Pre–processing

Pre–processing of image signal is the process where the image of subject in motion can be modified. This step is in a great measure affected by methodology of motion detection used in motion analysis. Main pre–processing methods include brightness transformations, geometric transformations, filtration, image sharpening, mathematical morphology etc.

To have the most effective marker–free analysis, the video record should be captured in the way no additional modifications are needed. Let us assume that the quality of acquired image is sufficient regarding possibilities of common video camera and that its digitized form will be convenient for further processing. Any arrangements or modifications are ineligible because our algorithm is based on searching of image areas similarity in consecutive frames. We suppose that the distortion in one frame will be similar to the distortion in the next one. The the algorithm is able to find it and pre–processing will not increase accuracy of motion detection.

5 Model Adaptation

5.1 Models

Simplified models of human body, designed according to the particular study by clinicians are adopted to the patients captured in video records. This is necessary to specify patient's position in one of the frames and for next automated motion tracking. Models are created as points, representing anatomical landmarks on patients bodies, that are interconnected by lines, usually representing body segments.

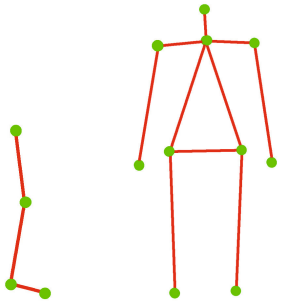


Fig. 5 Model of lower extremity specified by hip, knee and ankle used to analyse human gait in sagittal plane (left) and model of human body specified by forehead, neck, shoulders, wrists, pelvis and tip–toes used to analyse posture stability in frontal plane (right)

5.2 System Initialization

Ideally, the system should adopt model to the patient by itself. However, it is not possible to find an universal method to determine points of interest for different patients performing various movement activities, captured in various environments from different distances, using different video cameras etc. To allow system recognizing which part of the image represents particular body segments it easier to initiate them manually. The operator has just to select appropriate frame (it can be any in the sequence) and designate points according to the model so they are specified through the positions of individual picture elements.

In the case of sagittal plane, the hip joint can be given by coordinates (x_h, y_h), knee joint by (x_k, y_k), ankle by (x_a, y_a) and fifth metatarsal head by (x_m, y_m). Principle of this definition in the frame with the size of $L_1 \times L_2$ picture elements is shown in Fig. 6. Here, the individual points of this model can not be determined by single picture element because our method is based on searching of similar objects and it can be assumed that there will be several almost identical picture elements surrounding particular points of interest. Finally, using single picture element to define point of the model will cause the method will not be able to find correct position of these points in others frames of the video record.

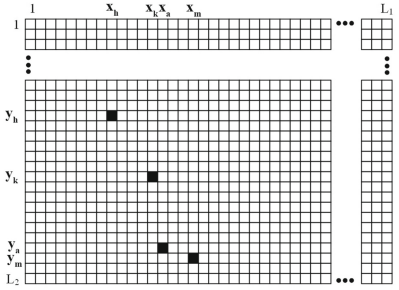


Fig. 6 Example of anatomical landmarks definition according to the lower extremity model in one of the frames with the size of the $L_1 \times L_2$ picture elements

To prevent above mentioned situation, every point of the model is specified by the matrix of surrounding picture elements in the size of $K \times K$, where K is positive odd integer. Then the matrix $\underline{J}$ surrounding point $X(x,y)$ for $K = 5$ is

$$\underline{J} = \begin{pmatrix} X(x-2,y-2) & X(x-1,y-2) & X(x,y-2) & X(x+1,y-2) & X(x+2,y-2) \\ X(x-2,y-1) & X(x-1,y-1) & X(x,y-1) & X(x+1,y-1) & X(x+2,y-1) \\ X(x-2,y) & X(x-1,y) & X(x,y) & X(x+1,y) & X(x+2,y) \\ X(x-2,y+1) & X(x-1,y+1) & X(x,y+1) & X(x+1,y+1) & X(x+2,y+1) \\ X(x-2,y+2) & X(x-1,y+2) & X(x,y+2) & X(x+1,y+2) & X(x+2,y+2) \end{pmatrix}, \quad (18)$$

where $X(x,y)$ is picture element representing point of the model located in the position x,y of current frame with dimensions $L_1 \times L_2$.

In general, small value of K reduces computing demands, but it may lead to decreased accuracy. Higher values of K increase number of calculations as well as accuracy of motion trajectory reconstruction. An optimal value should be find to keep small number of calculations and good reliability of reconstruction. Our experiments showed that the best ratio of these two requirements is obtained for values ranging from 9 to 13. Fig. 7 shows an example of ankle joint definition.

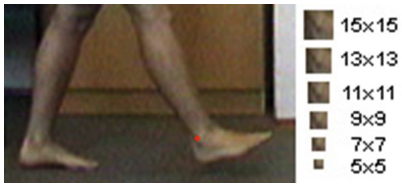


Fig. 7 Part of the frame capturing lower extremity in sagittal plane during gait and matrixes of picture elements specifying ankle for K=5, 7, 9, 11, 13 and 15

5.3 Calibration

In general, any motion analysis system should be calibrated to ensure measurement accuracy. Here, the calibration refers to the location of the video camera, usually capturing motion activities in frontal, sagittal or transversal plane and to the conversion of picture elements' distance on distance measured in units of length (meters). This is realized using known anthropometric measures taken from subjects' bodies and registered in their reports before analysis.

Let us assume that the calf length was measured as distance from knee joint to ankle joint and is defined as dc. Next, assume that the frame F captures also the knee joint K whose rotation axis in sagittal pane is given by the position of picture element (x_k, y_k) and in the case of ankle (A) it is element (x_a, y_a) as it is shown in Fig. 8.

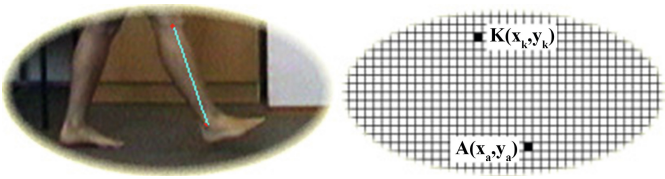


Fig. 8 Calf length specified as distance between centre of rotation in knee joint and centre of rotation in ankle joint

If the axis of video camera was perpendicular to sagittal plane, in which motion of the calf is realised, then the distance between picture elements specifying knee and ankle joints $|KA|$ is proportional to to the claf length dc.

$$d_p \leftrightarrow |KA|\,, \tag{19}$$

where dp is the distance in meters, and the distance of picture elements $|KA|$ is calculated as

$$|KA| = \sqrt{(x_k - x_a)^2 + (y_k - y_a)^2}. \tag{20}$$

The ratio of these two parameters gives information how many distance units falls to one picture element. Basically, it is possible to specify MPP parameter (meters per pixel) that is calculated as ratio

$$MPP = \frac{d_p}{|KA|}, [m/px]\,. \tag{21}$$

Value of MPP will be used to calculate kinematical parameters. Similarly, it can be derived for all segments of human body, whose main motion is realized in particular plane. Then the mean value of MPP will be given as

$$MMPP = \frac{1}{n}\sum_{i=1}^{n} MPP_i, \tag{22}$$

where MPP_i is the value for segment i and n is segment count. Error that arose from this calculation depends on the size and amount of picture elements in the frame. It's maximal value should be then defined as

$$\delta_{max} = \pm\frac{1}{2}px, \tag{23}$$

where px is the size of picture element. The advantage is that there in no need to calibrate camera before required motion is captured. Also, it is no matter what is the distance between camera and subject as MPP is recalculated for each record and for each subject during image processing phase.

6 Automatic Motion Tracking

Marker–free motion analysis tracks positions of anatomical landmarks that were specified using above mentioned method across the video sequence. Tracking process is finished when the positions of all model's points are derived in all frames of video record, i.e. trajectories are reconstructed. To do this in an automated way we used statistical properties of images. In general, the video signal has high intra and inter frame correlation and individual data depend on their surroundings.

The frame, in which the new position of the model will be searched will be considered as current one. Picture element from previous frame P_{F-1} will be used to predict position of picture element in current frame P_F with equal spatial coordinates (x,y). Because of motion occurred during the time between these two consecutive frames a motion compensation should be applied as it is shown in Fig. 9.

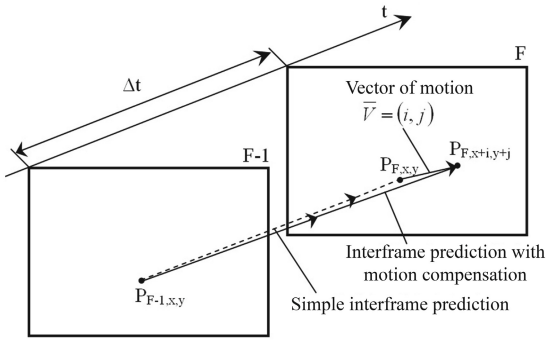


Fig. 9 Principle of interframe prediction with motion compensation

Assume that the element $P_{F-1,x,y}$ will move during time Δt to the point $P_{F,x+i,y+j}$. Then the motion of this point between two consecutive frames ($F-1$ and F) is defined by the motion vector $\overline{V} = (i, j)$.

Here, the vector of motion will be estimated using adaptation of blocks of picture elements. In this estimation the frames are subdivided into square blocks of picture elements with the size of $K \times K$ and the correlation between these blocks located in two consecutive frames is calculated. The correlation maximum gives position of motion vector. To increase computation rate a small searching area can be specified in current frame in which we expect to find motion vector as it can be seen in Fig. 10.

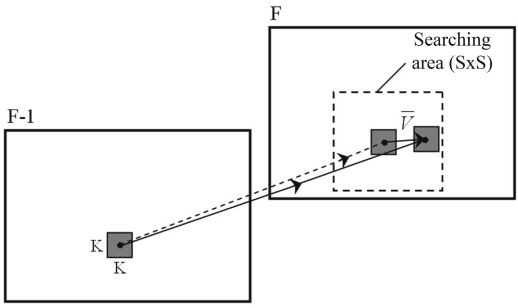


Fig. 10 Motion of block of picture elements between two frames in searching area given by motion vector

Let us assume that there is block of picture elements $\underline{J}_{F-1}$ defining one of the points of the model in the frame *F-1* and block of picture elements $\underline{J}_F$ from searching area of current frame F. Then the two dimensional correlation function will be calculated according to the following formula

$$R(i,j) = \sum_{m=1}^{K} \sum_{n=1}^{K} J_{F-1}(m,n) J_F(m+i, n+j), \tag{24}$$

where $J_{F-1}(m,n)$ and $J_F(m,n)$ are values of picture elements in previous and current frame, m and n are coordinates of picture element in particular block where

$$\begin{aligned} m &= 1,2,\ldots K, \\ n &= 1,2,\ldots K, \end{aligned} \tag{25}$$

K is the size of block of picture elements characterizing surrounding area of model's point and i and j are coordinates of block of picture elements displacement in current frame drawn in Fig. 11.

In order to be able to compare blocks of picture elements with various size and colour composition, it is necessary to normalize correlation function (24). Normalized two dimensional correlation function can be calculated as

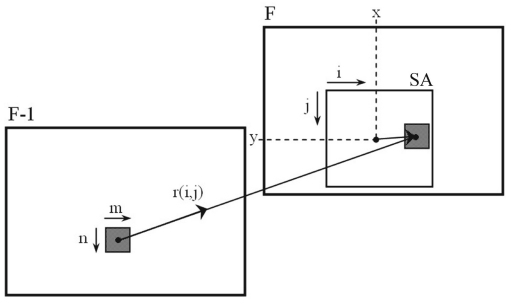


Fig. 11 Searching for correlation extreme between two blocks of picture elements from two consecutive frames

$$r(i,j) = \frac{\sum_{m=1}^{K}\sum_{n=1}^{K} J_{F-1}(m,n)\,J_F(m+i,n+j)}{\left[\sum_{m=1}^{K}\sum_{n=1}^{K} J_{F-1}^2(m,n)\right]^{\frac{1}{2}}\left[\sum_{m=1}^{K}\sum_{n=1}^{K} J_F^2(m+i,n+j)\right]^{\frac{1}{2}}}. \tag{26}$$

Then for motion vector applies

$$\overline{V} = \left\{\left(\widehat{i},\widehat{j}\right) : r\left(\widehat{i},\widehat{j}\right) \geq r(i,j)\,;\forall i,j \in SA\right\}, \tag{27}$$

where SA is searching area delimited by maximal displacement d_m of block of picture elements

$$\begin{aligned} -d_m \leq i \leq d_m,\\ -d_m \leq j \leq d_m. \end{aligned} \tag{28}$$

Searching area represents a part of frame where it is expected to identify movement of the subject. The size of searching area is selected with regard to velocity of motion in the scene. If the motion in the scene is small, then the potion between two consecutive frames will be also small and the searching area can be smaller. In the case of fast motions, higher information change is registered between two frames and the searching area should be bigger. An example of searching area is shown in following figure.

In general, the searching area has the size of $S \times S$ picture elements, while S is odd positive integer

$$S = K + 2d_m, \tag{29}$$

where K is the size of matrix defining surrounding area of the point of interest and d_m is the maximal expected displacement of the block of picture elements occurred between two frames. Centre of searching area is then given by the position of picture element (x,y) which specifies centre of particular point of the model defined in previous frame. Whole searching area covers all possible directions of motion. The main advantage is that the system is fully automated as there is no need to specify direction of motion in the scene.

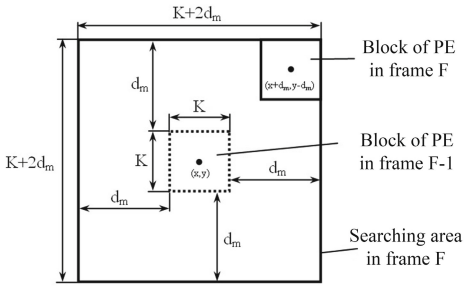


Fig. 12 Searching area used to find vector of motion in current frame F

The correlation is calculated for every picture element from searching area with block of picture elements specifying particular point of the model in previous frame. Then the result gives us matrix of correlation coefficients $\underline{R}$ with the size of $q \times q$.

$$\underline{R_j} = \begin{pmatrix} r(1,1) & r(1,2) & \cdots & r(1,q-1) & r(1,q) \\ r(2,1) & r(2,2) & \cdots & r(2,q-1) & r(2,q) \\ \vdots & & & & \vdots \\ r(q-1,1) & r(q-1,2) & \cdots & r(q-1,q-1) & r(q-1,q) \\ r(q,1) & r(q,2) & \cdots & r(q,q-1) & r(q,q) \end{pmatrix}, \tag{30}$$

where subject index j refers to the point of model, for which this matrix of correlation coefficients is calculated and q is given by the size of maximal possible displacement of the block of picture elements in all directions

$$q = 2d_m + 1. \tag{31}$$

Motion vector is then specified by the maximum value in the matrix of correlation coefficients according to the equation (27). Fig. 13 shows reconstructed trajectories of lower extremity's points.

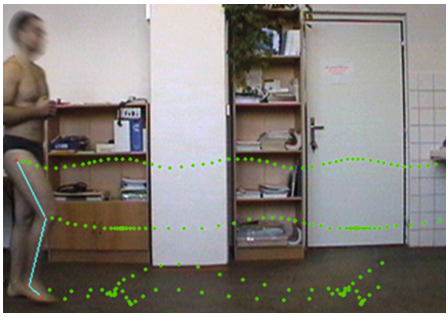


Fig. 13 Motion trajectories of lower extremity model reconstructed by automated motion tracking algorithm in gait analysis

The mechanism of automatic motion tracking algorithm can be summarized using following list of steps that has to be realized across video sequence.

1. Enter the size of searching area $S \times S$ and clear correlation coefficients matrix.
2. Load the block of picture elements ($K \times K$) specifying point of the model in previous frame.
3. Load the block of picture elements ($K \times K$) from searching area of current frame.
4. Calculate correlation of block of picture elements from searching area and block of picture elements specifying point of the model in previous frame.
5. Repeat step 3 and 4 for all blocks of picture elements in searching area.
6. Motion vector and position of the model's point in current frame are given by correlation maximum.
7. If there is a need to track another point of the model, then go to step 1 otherwise finish motion tracking in current frame.
8. If there is a eed to track motion in the next frame, then go to the next frame and continue with the step 1. Otherwise finish motion tracking in the video sequence.

7 Kinematical Parameters

Motion trajectories specified by space coordinates and *MPP* ratio are used to derive both the anatomical joint angles and time–spatial parameters. These will be essential to evaluate quality of analysed motion activities as well as to identify potential pathologies in patients through comparison with normal motion patterns that are usually obtained from healthy subjects. Parameters used to verify the proposed method functionality relate to gait analysis and include

- trajectories of motion - positions, velocities and accelerations of individual joints,
- anatomical joint angles - hip flexion/extension, knee flexion/extension, and ankle plantar/dorsal flexion,
- spatio–temporal parameters - gait cycle length, gait cycle time, gait cycle velocity, cadence (cycles per minute), stance phase and swing phase.

Time values were derived from the frame rate

$$t = \frac{p-1}{f}, \tag{32}$$

where f is the frequency of frames and p is the number of the frame in video sequence for which the time is calculated. Then the time between two consecutive frames is given be the ratio $1/f$.

Averaged velocities are calculated as ratio of changes in distance and time

$$v = \frac{\Delta s}{\Delta t}, \tag{33}$$

and averaged acceleration of particular point is calculated as second derivation of trajectory of as first derivation of velocity. If the velocity is known, then the acceleration is given as

$$a = \frac{\Delta v}{\Delta t}, \tag{34}$$

Anatomical joint angles in particular plane (sagittal, frontal, transversal) reflect anatomical structures and directions. One of the ways to calculate these plane angles is to calculate body segments dimensions and apply them in cosine theorem

$$\alpha = arccos \frac{a^2 + b^2 - c^2}{2ab}, \tag{35}$$

where a, b and c are sides of triangle in particular joint given by the model as it is shown in Fig. 14. Calculations of all these kinematical parameters were integrated together with motion tracking algorithms into marker–free motion analysis system.

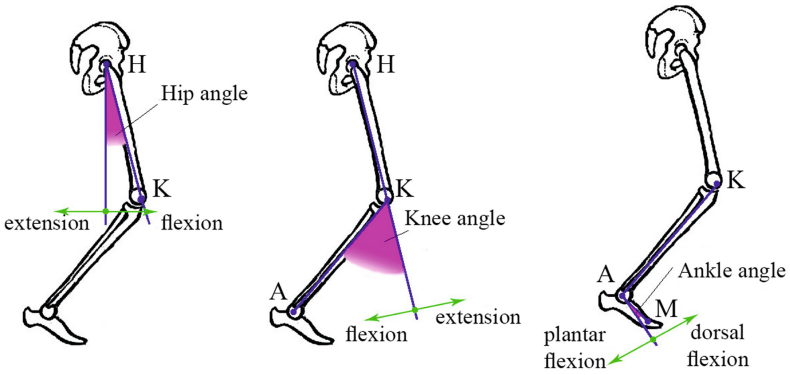


Fig. 14 Anatomical joint angles of lower extremity joints – hip, knee and ankle specified in sagittal plane

The system was tested on group of ten young healthy adult male subjects in the average age of 24.8 ± 2.18 years and average height 178.54 ± 5.4 cm. The subjects had no known pathology and they were asked to walk at their natural speed during the laboratory tests. The data of these healthy subjects were also used to create normative gait database, which is usually used by gait laboratories in clinical trials to identify pathologies in patients with motion disorders. The normative gait database of three anatomical joint angles (hip flexion/extension, knee flexion/extension and ankle plantar/dorsal flexion angles) is shown in Fig. 15 .

Here, the anatomical joint angles were standardised to one gait cycle, where the data of all subjects were defined in identical points of gait cycle. Cubic spline interpolation was used in all angles and for all included subject to do it. Definition of normative gait data was made by determination of upper and lower bounds of reliability interval. As for the references in literature and the most used values in mathematical statistics and laboratories we used the probability of 95% reliability interval coverage. This interval means that the data included in are the data of normal healthy subjects and the data out of this interval represent certain type of pathological pattern in subjects gait.

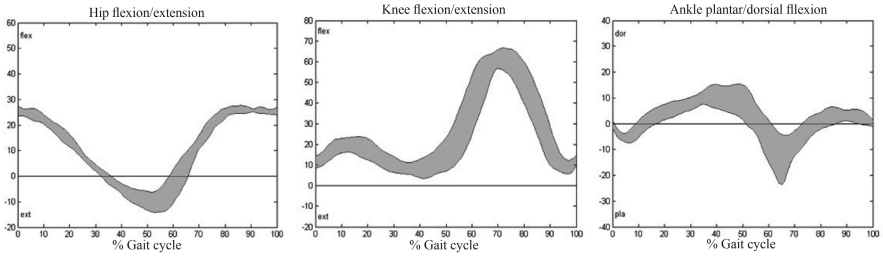


Fig. 15 Normative gait database of anatomical joint angles for the hip, knee and ankle in stagittal plane

8 Conclusion

Newly developed marker–free analysis method was described in this chapter. The algorithms were implemented into the motion analysis system, whose functionalities were verified at the group of young healthy male subjects. Here, the trajectories of motion, and kinematics parameters were acquired and normative gait database was generated. The experimental results convinced us that the designed methodology and the designed system are clinically applicable and are not limited to only laboratory usage. Determination of normative gait database might help to utilize the system in the process of pathological gait assessment.

Acknowledgements. This work has been elaborated in the framework of the project KEGA 005UPJS–4/2012 (50%) and in the framework of the project "Support research and development in the Moravian–Silesian Region 2013 DT 1 – International research teams" (RRC/05/2013), financed from the budget of the Moravian–Silesian Region (50%).

References

1. Pediaditis, M., Tsiknakis, M., Leitgeb, N.: Vision-based motion detection, analysis and recognition of epileptic seizures A systematic review. Computer Methods and Programs in Biomedicine 108, 1133–1148 (2012)
2. Porada, V., et al.: Identifikace osob podle dynamického stereotypu chůze, Karlovy Vary, Vysoká škola Karlovy Vary (2010) ISBN 978-80-87236-01-7
3. Courtney, J., de Paor, A.M.: Monocular Marker-Free Gait Measurement System. IEEE Transactions on Neural Systems and Rehabilitation Engineering 18(4), 453–460 (2010)
4. Penhaker, M., Kasik, V., Snasel, V.: Biomedical distributed signal processing and analysis. In: Saeed, K., Chaki, R., Cortesi, A., Wierzchoń, S. (eds.) CISIM 2013. LNCS, vol. 8104, pp. 88–95. Springer, Heidelberg (2013)
5. Sholukha, V., Bonnechere, B., Salvia, P., Moiseev, F., Rooze, M., Van Sint Jan, S.: Model-based approach for human kinematics reconstruction from markerless and marker-based motion analysis systems. Journal of Biomechanics 46(14), 2363–2371 (2013)
6. Macurova, A., Macura, D.: Asymptotic properties of the variables of the roughness surface, Applied Surface Science (2012), DOI: 10.1016/j.apsusc.2012.10.086

7. Živčák, J., Hudák, R., Tóth, T.: Rat skin wounds tensile strength measurements in the process of healing, SAMI, Article number 6208996, pp. 389-392 (2012)
8. Majerník, J.: Marker-free Method for Reconstruction of Human Motion Trajectories. Advances in Optoelectronic Materials 1(1), 4–9 (2013)
9. Han, R., Zhang, F., Wan, X., Fernndez, J., Sun, F., Liu, Z.: A marker-free automatic alignment method based on scale-invariant features. J. of Struct. Biology 186(1), 167–180 (2014)
10. Ota, S., et al.: A comparison of knee joint biomechanics during free gait and cartilage T2 mapping values in healthy individuals in their twenties and forties. Osteoarthritis and Cartilage 124, S123–S124 (2014)
11. Zhang, J.: Basic gait analysis based on continuous wave radar. Gait & Posture 36(4), 667–671 (2012)
12. Penhaker, M., Krejcar, O., Kasik, V., Snášel, V.: Cloud Computing Environments for Biomedical Data Services. In: Yin, H., Costa, J.A.F., Barreto, G. (eds.) IDEAL 2012. LNCS, vol. 7435, pp. 336–343. Springer, Heidelberg (2012)
13. Mihalík, J., Štefanišin, R., Michalčin, V.: Medzisnímkové hybridné kódovanie videosignálov. Nové smery v spracovaní signálov, L, 223–226 (2002) ISBN 8080401802

Part II
Computational Intelligence and Biology

Process Calculus and Illocutionary Logic for Analyzing the Behavior of Schistosomatidae (Trematoda: Digenea)

Andrew Schumann and Ludmila Akimova

Abstract. The behavior of collectives of trematode larvae (miracidia and cercariae) can be considered adaptable to environmental changes. For describing that behavior, we appeal to a language which is a kind of process calculus called *Physarum* spatial logic that is based on a formalization of the behavior of *Physarum polycephalum*, the supergroup Amoebozoa, phylum Mycetozoa, and class Myxogastria. This logic shows in what way we can program the behavior of *Physarum polycephalum* by attractants and repellents. This language contains labels for attractants and repellents. The behavior of miracidia and cercariae collectives can be simulated within this language as well. Taking into account the fact that the behavior of miracidia and cercariae collectives can be programmed only by attractants (repellents for miracidia and cercariae are not known still), we can claim that the behavior of miracidia and cercariae is a restricted (poorer) form of *Physarum* spatial logic. Within *Physarum* spatial logic we can define illocutionary logic showing how intelligent in the human meaning the behavior coded in *Physarum* spatial logic is. Thus, this logic can show how intelligent the behavior of trematode larvae (miracidia and cercariae) collectives is. So, illocutionary logic allows us to apply the Turing test to *Physarum polycephalum* as well as to trematode larvae to show their ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human.

Andrew Schumann
Department of Cognitivistics,
University of Information Technology and Management in Rzeszow,
Rzeszow, Poland
e-mail: Andrew.Schumann@gmail.com

Ludmila Akimova
State Scientific and Production Amalgamation,
The Scientific and Practical Center of the National Academy of Sciences
of Belarus for Bioresources, Minsk, Belarus
e-mail: akimova_minsk@mail.ru

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*,
Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_5

1 Introduction

Now we are working on the programmable amorphous biological computer based on plasmodia of *Physarum polycephalum*. This abstract computer is called *slime mould based computer*. The plasmodium behaves and moves as a giant amoeba and its behavior can be considered a biological implementation of Kolmogorov-Uspensky machines [2, 3]. This allows us to use the plasmodia of *Physarum polycephalum* for solving different computing tasks that can be solved in Kolmogorov-Uspensky machines as well. The slime mould based computer is programmed using attractants and repellents. On the one hand, it was experimentally proved that the slime mould prefers substances with potentially high nutritional value, e.g. it is attracted by peptones, aminoacids phenylalanine, leucine, serine, asparagine, glycine, alanine, aspartate, glutamate, and threonine. On the other hand, repellents for *Physarum polycephalum* can be presented by some illumination-, thermo- and salt-based conditions. Usually, the plasmodium forms a congregation of protoplasm in food sources to surround them, secret enzymes, and digest the food. The slime mould based computer [2] can be regarded as a parallel computing substrate complementary to existing massive-parallel reaction-diffusion chemical processors [1].

In papers [28, 29], we have showed that the behavior of plasmodia of *Physarum polycephalum* has an own spatial logic, which is one of the natural implementations of process calculus. This logic called *Physarum spatial logic* can be used as a programming language for the slime mould based computer. Taking into account the fact that within process calculus we can formalize and describe different concurrent processes, within *Physarum* spatial logic we can do the same as well.

In this paper, we will show that the behavior of collectives of the genus *Trichobilharzia* Skrjabin & Zakharov, 1920 (Schistosomatidae Stiles & Hassall, 1898) can be simulated by *Physarum* logic. This means that, first, a local group of Schistosomatidae can behave as a programmable biological computer, second, a biologized kind of process calculus such as *Physarum* spatial logic can describe concurrent biological processes at all.

On the one hand, plasmodia can be used in approximating a range of path planning, computational geometry, combinatorial optimization, and robotics problems [2]. It means that the behavior of trematode larvae can be used in the same approximations, too. On the other hand, plasmodia can be considered a programmable medium for simulating emergent transport networks and some other intelligent processes of human activity [5, 20]. Therefore we can ask the question, how intelligent in the human meaning the behavior coded in *Physarum* spatial logic is. Respectively, how intelligent the behavior of trematode larvae (miracidia and cercariae) collectives is. In other words, we can try to check a version of the Turing test showing abilities of *Physarum polycephalum* and trematode larvae to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. The version we will use is called the *illocutionary Turing test*.

What does the term 'illocutionary' mean? It means to be the communicative effect (as commanding or requesting) of an utterance. These effects are studied in *lllocutionary logic* that explains, how context can affect the meaning of speech acts.

The first formalization of illocutionary logic was created by J.R. Searle and D. Vanderveken [30]. Recall that in classical logic a well-formed proposition is evaluated as either true or false and in conventional logics as a degree of truth. However, speech acts that contain a speaker's illocutionary forces such as 'I order you to leave the room,' 'I am thinking about' are evaluated as *either successful or unsuccessful* in the given context of utterance (notice that within the more detailed consideration in the same way as in non-classical logics, the meaning of non-defective simple performative proposition could be evaluated as a degree of successfulness). Thus, the successfulness of speech acts depends on human actions and utterance contexts. In other words, the human behavior is a semantics for speech acts. The *illocutionary Turing test means checking, if the behavior of a person or machine can play a role of semantics for some speech acts*, i.e., on the one hand, if the behavior of that person or machine can be understood within human interactions and, on the other hand, if that behavior can verify some successful speech acts.

Speech acts from the logical point of view are said to be *illocutionary propositions*. Whenever a speaker utters a sentence in an appropriate context with certain intentions, he performs one or more illocutionary propositions. Let us show that the process calculus (e.g. *Physarum* spatial logic) can play the role of *formal* semantics for illocutionary propositions. This semantics contain formal descriptions of processes, which can have different implementations in reality. Let us consider the proposition '*I'll buy this thing.*' I can do it. This proposition is illocutionary and consists of the following: the one informative proposition describing something that I am going to buy and the second performative proposition appealing to my promise to buy. This promise means that the proposition '*I'll buy this thing*' is true at an actual context of utterance denoted by $x \in W$ if for any process, when I am buying something if there is a transition $x \xrightarrow{a} y$, where a from Act, the set of action of process calculus I can use for verifying, is my action to buy, then '*this thing*' I am going to buy is true at the future context of utterance $y \in W$ (i.e. '*this thing*' is true in each future situation, starting from this actual world due to a). Consequently, we can understand illocutionary propositions as modal: $\bigcirc_a \Psi$:= '*I swear to buy* Ψ' (as strong promise) and $\diamondsuit_a \Psi$:= '*perhaps I would like to buy* Ψ' (as weak promise):

$x \models \bigcirc_a \Psi$ if for any process P containing a (i.e. my action to buy), we have that if P contains a transition $x \xrightarrow{a} y$, then $y \models \Psi$.

$$\diamondsuit_a \Psi := (\bigcirc_a (\Psi \Rightarrow \bot) \Rightarrow \bot).$$

The same semantics takes place if a is not only permission (the particular case of a comissive performative verb [30]), but in all cases where a is a comissive performative verb and directive performative verb. In a case where a is an assertive performative verb, declarative performative verb, or expressive performative verb [30] we have the following semantics of illocutionary propositions $\bigcirc_a \Psi$, $\diamondsuit_a \Psi$:

$y \models \bigcirc_a \Psi$ if for any process P containing a (e.g. my action to buy), we have that if $x \models \Psi$, then P contains a transition $x \xrightarrow{a} y$.

$\Diamond_a \Psi := (\bigcirc_a (\Psi \Rightarrow \bot) \Rightarrow \bot)$.

For example, if $\Diamond_a \Psi$ '*maybe I would like to buy* Ψ' has a declarative illocutionary force (not comissive), then its semantics is as follows:

$y \models \Diamond_a \Psi$ if for some processes containing a we have that $x \xrightarrow{a} y$ and not $x \models \Psi$.

Thus, illocutionary propositions are realized not on static databases, but on sets of processes. Then we can apply the illocutionary Turing test on real processes coded in process calculus, if they can verify some illocutionary acts. Let us show that the activity of an organization can be described in terms of process calculus, i.e. this activity can seem intelligent for us, i.e. they can satisfy the illocutionary Turing test. For example, it is possible to construct a process automaton for modelling processes *call*, *write*, *meet*:

W_{call} = {'call a subordinate,' 'call a client,' 'call a manager'}.

W_{write} = {'write a report for a client with data collected by a subordinate,' 'write a general report for a manager'}.

W_{meet} = {'meet a client'}.

What does the illocutionary proposition '*I'll surely call my boss*' mean, if the organization activity combines different processes *call*, *write*, *meet* simultaneously? I cannot use just a recursion *call* = ⟨*call my boss*, *call*⟩. In situations of complex processes the meaning of $\bigcirc_{call} \Psi$ '*I'll surely call my boss*' and $\bigcirc_{call} \Psi$ '*Maybe I'll call my boss*' will be understood as follows:

$x \models \bigcirc_{call} \Psi$ if for any process P, such that $W_P = W_{call} \cup W_{write} \cup W_{meet}$, we have that if P contains a transition $x \xrightarrow{a} y$, then $y \models \Psi$, where *call* is my calling.

Thus, for defining meanings of illocutionary propositions we can focus on streams $xy\ldots$ of states in a process. Any activity analyzed in the frames of process calculus can be considered semantics of corresponding illocutionary propositions. In this paper, we will show a verification of the most primitive illocutionary propositions such as '*I would like to eat*' and '*I fear this thing*' (e.g. '*I fear the light*') in the activity of living organisms such as *Physarum polycephalum* and trematode larvae of Schistosomatidae. It means that they satisfy the illocutionary Turing test for some very simple speech acts.

2 *Physarum* Spatial Logic

In this section, let us consider some basics of *Physarum* spatial logic. The behavior of *Physarum polycephalum* plasmodia can be divided into the following elementary processes: inaction, fusion, cooperation, and choice, which could be interpreted as unconventional (spatial) falsity, conjunction, weak and strong disjunction respectively, denoted by *Nil*, &, |, and +. These operations differ from conventional ones, because they cannot have a denotational semantics in the standard way. However, they may be described as special spatial transitions over states of *Physarum polycephalum* machine: inaction (*Nil*) means that pseudopodium has just stopped to behave; fusion (&) means that two pseudopodia come in contact one with another and then merge; cooperation (|) means that two pseudopodia behave concurrently; choice (+) means a competition between two pseudopodia in their behaviors or their split.

Some explanations of basic operations of *Physarum* spatial logic are as follows:

Prefix $a.P$ – Let x be an event and P a process. Then $(x \xrightarrow{a} P)$ (it is read: 'P after x') describes an object which occurs first in event x, and then behaves exactly as P. This phenomenon is denoted by $a.P$.

Choice $P_1 + P_2$ – If x and y are different events, $x \xrightarrow{a} P_1$ and $y \xrightarrow{a} P_2$ do not have joint states, then we can obtain a new process $P_1 + P_2$ by adding a new state z, assuming that the process behaves further either as $x \xrightarrow{a} P_1$ or as $y \xrightarrow{a} P_2$. In other words, the subsequent behavior of object $P_1 + P_2$ is described by the process P_1, if an event x appears first, or by the process P_2, if an event y appears first.

Fusion $P_1 \& P_2$ – A fusion holds between events if they are simultaneously carried out. If P_1 and P_2 are processes with an identical alphabet, then $P_1 \& P_2$ designates a process behaving as system, combining processes P_1 and P_2 and the fusion between them is synchronized step by step. For example, a chocolate can be purchased, only when it is wanted by a customer and a shop assistant is ready to sell it. The fusion $P_1 \& P_2$ in some cases can be described as mutual recursion. For example, let us consider the process of making authorization on documents. A subordinate and manager participate in this process. Subordinate's action consists in the compilation of document (*doc*) and its preparation for authorization (*authorization*). The manager's action consists in signing the document (*sign*) or its rejection (*reject*). The equations defining the alphabet and behavior of three processes are defined as follows: $W_P = W_a = W_b = \{doc, authorization, sign, reject\}$; $P = (doc \to a | doc \to b)$; $a = (authorization \to sign \to a | authorization \to b)$; $b = (authorization \to reject \to b \ | authorization \to a)$. We see that $P = a \& b$. In other cases, $P_1 \& P_2$ can appear as a new process which has arisen as a result of fusion of P_1 and P_2. If x and y are different events, $x \xrightarrow{a} P_1$ and $y \xrightarrow{a} P_2$ do not have joint states, then we can obtain a new process $P_1 \& P_2$, such that $x \xrightarrow{a} P_1 \& P_2$ and $y \xrightarrow{a} P_1 \& P_2$.

Concurrency $P_1 | P_2$ – If x and y are different events, $x \xrightarrow{a} P_1$ and $y \xrightarrow{a} P_2$ do not have joint states, then we can obtain a new process $P_1 | P_2$, such that in the same point

of time actions can be fulfilled by the processes P_1, P_2, so that a new process $P_1|P_2$ can simultaneously commit couples of actions. Formally:

1. for any transition $x_0 \xrightarrow{a} x_1$ from P_1 and for any state y from P_2 the process $P_1|P_2$ contains a transition $\langle x_0, y\rangle \xrightarrow{a} \langle x_1, y\rangle$;
2. for any transition $y_0 \xrightarrow{a} y_1$ from P_2 and for any state x from P_1, the process $P_1|P_2$ contains a transition $\langle x, y_0\rangle \xrightarrow{a} \langle x, y_1\rangle$;
3. for any couple of transitions $x_0 \xrightarrow{a} x_1$ from P_1 and $y_0 \xrightarrow{\overline{a}} y_1$ from P_2 with complementary actions a and $\overline{a}$, the process $P_1|P_2$ contains a transition $\langle x_0, y_0\rangle \xrightarrow{\tau} \langle x_1, y_1\rangle$.

For example, a parallel process takes place, if a manager accepts several documents before signing or rejecting them. Another example is presented by a worker performing several jobs and between jobs there can be a concurrency.

Hiding $P_1 \setminus P_2$ – On the basis of process P_1 we can obtain a process $P_1 \setminus P_2$ by means of deleting those transitions which are obtained by means of labels with names from P_2. Such hiding does not concern interior labels τ.

Thus, a process calculus for describing the dynamics of *Physarum polycephalum* machine is presented as a labelled transition system with some logical relations.

Assume that there are $N = 1, 2, \ldots, i, \ldots$ active species or growing pseudopodia and a state of one growing pseudopodium i is denoted by $p_i \in \mathscr{L}$. These states are time dependent and they are changed by different stimuli (actions) from T. The transitions from one state to another are defined as $\{p \xrightarrow{a} q\colon\ p, q \in \mathscr{L}, a \in T\}$. So, each action $a \in T$ starts on a state p, which is its current position, and says about a transition (propagation) of a state p to another state q. There are the following stimuli (actions) on growing pseudopodia:

- The set of *attractants* $\{A_1, A_2, \ldots\} \subset T$. They are sources of nutrients, on which the plasmodium feeds. Each attractant A is a function from $\mathscr{L}$ to $\mathscr{L}$.
- The set of *repellents* $\{R_1, R_2, \ldots\} \subset T$. Plasmodium of *Physarum polycephalum* avoids light and some thermo- and salt-based conditions. Each repellent R is a function from $\mathscr{L}$ to $\mathscr{L}$.
- The set of *diffusions* (new active zones of growing pseudopodia) $\{C_1, C_2, \ldots\} \subset T$. $C(p)$ means a diffusion of growing pseudopodia $p \in \mathscr{L}$.

Hence, $T = \{A_1, A_2, \ldots\} \cup \{R_1, R_2, \ldots\} \cup \{C_1, C_2, \ldots\}$.

Our process calculus contains the following basic operators: *Nil* (inaction), . (prefix), | (cooperation), \ (hiding), & (reaction/fusion), + (choice), a (constant or restriction to a stable state), $A(\cdot)$ (attraction), $R(\cdot)$ (repelling), $C(\cdot)$ (spreading/diffusion). Let $T = \{a, b, \ldots\}$, the set of all actions (evidently, this set is finite), be considered as a set of names. A name refers to a communication link or channel. With every $a \in T$ we associate a complementary action $\overline{a}$. Let us suppose that a designates an input port and $\overline{a}$ designates an output port. Any behavior of *Physarum polycephalum* will be considered outputs and any form of outside control and stimuli will be considered appropriate inputs. So, T contains ports that can be interpreted under different conditions as input ports or output ports. So, for each

$X \in \{A_1, A_2, \ldots\} \cup \{R_1, R_2, \ldots\}$, we take that $X(p) = q$ is a function from $\mathscr{L}$ to $\mathscr{L}$ such that X is a stimulus for p and X makes an output $q \in \mathscr{L}$. Let $X(p)$ denote an input and $\overline{X(p)}$ an output.

Define $\mathbf{L}$ be the set of labels built on T (under this interpretation, $a = \overline{\overline{a}}$). Suppose that an action a communicates with its complement $\overline{a}$ to produce the internal action τ and τ belongs to $\mathbf{L}$, too.

We use now the symbols γ, δ, ..., etc., to range over labels (actions), with $a = \overline{\overline{a}}$, and the symbols P, Q, etc., to range over processes on states p_i. The processes are given by the syntax:

$$P, Q := Nil \mid \gamma.P \mid A(p).P \mid \overline{A(p)}.P \mid R(p).P \mid \overline{R(p)}.P \mid C(p).P \mid (P|Q) \mid P \setminus Q \mid P\&Q \mid P + Q \mid a$$

Each label is a process, but not vice versa, because a process may consists of many labels combined by the basic operators.

An operational semantics for this syntax is defined as follows:

$$\text{Prefix}: \quad \frac{}{\gamma.P \xrightarrow{\gamma} P},$$

$$\frac{}{A(p).P \xrightarrow{\delta} P} \ (A(p) = \delta), \quad \frac{}{\overline{A(p)}.P \xrightarrow{\delta} P} \ (\overline{A(p)} = \delta),$$

$$\frac{}{R(p).P \xrightarrow{\delta} P} \ (R(p) = \delta), \quad \frac{}{\overline{R(p)}.P \xrightarrow{\delta} P} \ (\overline{R(p)} = \delta),$$

(the conclusion states that the process of the form $\gamma.P$ (resp. $A(p).P$, $\overline{A(p)}.P$, $R(p).P$, $\overline{R(p)}.P$) may engage in γ (resp. $A(p)$, $\overline{A(p)}$, $R(p)$, $\overline{R(p)}$) and thereafter they behave like P; in the presentations of behaviors as trees, $\gamma.P$ (resp. $A(p).P$, $\overline{A(p)}.P$, $R(p).P$, $\overline{R(p)}.P$) is understood as an edge with two nodes: γ (resp. $A(p)$, $\overline{A(p)}$, $R(p)$, $\overline{R(p)}$) and the first action of P),

$$\text{Diffusion}: \quad \frac{P \xrightarrow{\overline{A(p)}} P'}{P \xrightarrow{\overline{A(p)}} C(q)} \ (C(q) = P', \overline{A(p)} = q),$$

$$\text{Constant}: \quad \frac{P \xrightarrow{\gamma} P'}{a \xrightarrow{\gamma} P'} \ (a = P, a \in \mathbf{L}),$$

$$\text{Choice}: \quad \frac{P \xrightarrow{\gamma} P'}{P + Q \xrightarrow{\gamma} P'}, \quad \frac{Q \xrightarrow{\gamma} Q'}{P + Q \xrightarrow{\gamma} Q'},$$

(these both rules state that a system of the form $P + Q$ saves the transitions of its subsystems P and Q),

$$\text{Cooperation}: \quad \frac{P \xrightarrow{\gamma} P'}{P|Q \xrightarrow{\gamma} P'|Q}, \quad \frac{Q \xrightarrow{\gamma} Q'}{P|Q \xrightarrow{\gamma} P|Q'},$$

(according to these rules, the cooperation $|$interleaves the transitions of its subsystems),

$$\frac{P \xrightarrow{\gamma} P' \qquad Q \xrightarrow{\overline{\gamma}} Q'}{P|Q \xrightarrow{\tau} P'|Q'},$$

(i.e. subsystems may synchronize in the internal action τ on complementary actions γ and $\overline{\gamma}$),

$$\text{Hiding}: \quad \frac{P \xrightarrow{\gamma} P'}{P\backslash Q \xrightarrow{\gamma} P'\backslash Q} \quad (\gamma \notin Q, Q \subseteq \mathbf{L}),$$

(this rule allows actions not mentioned in Q to be performed by $P \setminus Q$),

$$\text{Fusion}: \quad \frac{}{\gamma.P\&\check{P} \xrightarrow{\gamma} Nil}$$

(the fusion of dual processes are to be performed into the inaction, e.g. a fusion of an attractant/repellent P and appropriate repellent/attractant $\check{P}$),

$$\frac{P \xrightarrow{\gamma} P' \quad Q \xrightarrow{\gamma} P'}{P\&Q \xrightarrow{\gamma} P'}, \qquad \frac{P \xrightarrow{\gamma} P' \quad Q \xrightarrow{\gamma} P'}{Q\&P \xrightarrow{\gamma} P'}$$

(this means that if we obtain the same result P' that is produced by the same action γ and evaluates from two different processes P and Q, then P' may be obtained by that action γ started from the fusion $P\&Q$ or $Q\&P$).

These are inference rules for basic operations. The ternary relation $P \xrightarrow{\gamma} P'$ means that the initial action P is capable of engaging in action γ and then behaving like P'.

Some basic properties of operations defined above:

$$Nil\backslash P \cong Nil,$$

$$P\&\check{P} \cong Nil,$$

$$P\&P \cong P,$$

$$P\&Nil \cong Nil,$$

$$(P+Q)\backslash P' \cong P\backslash P' + Q\backslash P',$$

$$(P\&Q)\backslash P' \cong P\backslash P'\&Q\backslash P',$$

$$P\&Q \cong Q\&P,$$

$$P\&(Q\&R) \cong (P\&Q)\&R,$$

$$P+P \cong P,$$

$$P+Nil \cong P,$$

$$P+Q \cong Q+P,$$

$$P+(Q+R) \cong (P+Q)+R,$$

$$P\&(Q+R) \cong (P\&Q)+(P\&R),$$
$$P+(Q\&R) \cong (P+Q)\&(P+R),$$

where $\cong$ is a congruence relation defined on the set of processes.

Now we can show that in the behavior of any local group of Schistosomatidae we can observe the same elementary processes: inaction, fusion, cooperation, and choice, which are defined in the same way.

3 Life Cycle of *Trichobilharzia szidati* (Diginea: Schistosomatidae)

First of all, let us consider the life cycle of *Trichobilharzia szidati*, one of the members of the family Schistosomatidae. It is a subclass Digenea Carus, 1863 (Platyhelminthes: Trematoda), which are exclusively endoparasites of animals. The digenean life cycle has the form of heterogony, i.e. there is a natural alternation of amphimictic (usually synarmophytous) and parthenogenetic stages. At these stages digeneans have different outward, different means of reproduction, and different adaptation to different hosts. The interchange of hosts is necessary for a successful realization of digenean flukes life cycle. The majority of representatives of this subclass have a complete life cycle with participation of three hosts: intermediate, additional (metacercarial) and final. Molluscs are always the first intermediate hosts, while different classes of vertebrate animals are final hosts.

Among digeneans, there are parasites, belonging to the family of Schistosomatidae, which represents have adapted to parasitizing in the circulatory system of vertebrate animals. Puberal representatives of this family are diecious individuals (in other digenean families maritas are hermaphrodites). The family includes the following three subfamilies: Schistosomatinae Stiles and Hassall, 1898, they parasitize a variety of birds and mammals, including human being; Bilharziellinae Price, 1929 and Gigantobilharziinae Mehra, 1940, they parasitize birds. Representatives of the first subfamily (in particular the genus *Schistosoma* Weinland, 1858) can parasitize humans. In the tropical and subtropical countries, about 200 million persons are infected by them from which 11 thousand persons annually die because of the given infestation [13]. Representatives of the latter two subfamilies, the so-called avian schistosomatidae, have been observed in all continents, including Europe. In a puberal state, they parasitize birds, however they are capable to incorporate into a human organism as nonspecific host. After they penetrate human skin, where they perish, they invoke allergic dermatitis. The fact of incorporation of these larvae into nonspecific hosts invokes an interest to avian schistosomatidae. Therefore the simulation of behavior of their local groups can be interesting from a medicine view, because it allows better to understand features of digenean behavior. Simulating their behavior is possible by means of *Physarum* spatial logic as we will show.

The life cycle of all representatives of the family Schistosomatidae is identical. Members of this family have two free-swimming dispersal stages (miracidia and

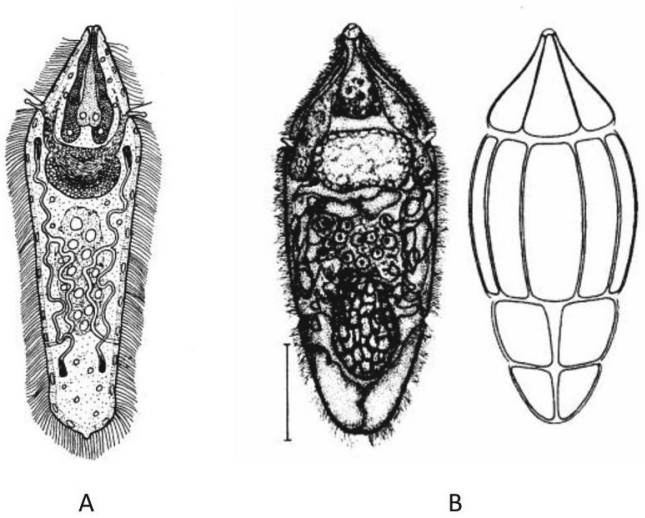


Fig. 1 The morphology of miracidia. A. the general morphology of *Trichobilharzia szidati* (from [26]); B. the general morphology of *Trichobilharzia regenti* (left) and the arrangement of ciliated plates (right), scale bar = 25mm (from [19])

cercariae) which actively search for their hosts (intermediate and final respectively). The successful search for the next owners is carried through behavior and host-specificity which elaborated during a long coevolution between schistosomatides and their hosts. Miracidial and cercarial host-finding is initiated mainly by response to gravitational, light and chemical stimuli.

Thus, the specificity and diversity of host-recognition strategies suggest that finding hosts and their recognition are important determinants in the evolution of parasite life cycles.

3.1 Miracidia (Genus Trichobilharzia)

3.1.1 Morphology of Miracidia

Miracidia are free-swimming larval stage of parthenogenetic generation of Schistosomatidae (figure 1). They hatch from eggs which got to water from organisms of final hosts. Adult schistosomatides produce eggs that contain fully developed miracidia.

Miracidia have an elongated body. The body surface is covered with four rows of epithelial plates that carry a multitude of cilia involved in active motions of miracidia (figure 1, B). The anterior part of the body contains terminal openings

of the following penetration unicellular glands: two cells of apical glands and two cells of lateral glands. The posterior part of the body contains clusters of germ cells from which daughter sporocyst is formed subsequently. Miracidia lack any digestive system and they cannot feed.

3.1.2 Behavior of Miracidia

Miracidia of *Trichobilharzia szidati* [26] hatch from eggs within a short time 5–10 min. In order to survive, miracidia must infect a snail for 20h at 24°C before they die [26]. This larval stage shows a specialized behavior that allows them to find a snail-host [14, 16]. The positive photo- and negative geotactic orientation is an adaptation to reach the preferred habitats of their host-snail species [14, 15, 17, 32, 33]. Also, it is now confirmed that miracidia respond to chemical host cues when they approach and invade their host snails. There are several studies to characterize the chemical nature of these host signals [23, 27, 31, 32]. Snail-hosts release various secreted/excreted products around them in water. Miracidia are attracted by the macromolecular components of these products, which consist of glycoproteins larger than 30kDa, termed "miracidia attracting glycoproteins" [21]. These glycoproteins produced by different species of snails have similar peptide-based, but differ by saccharide component. Miracidia can differentiate between these glycosylation patterns in order to find and infect mainly specific snail species [21, 22]. So, the chemical cues secreted by the freshwater gastropod *Lymnaea stagnalis* stimulate behavioral responses of *Trichobilharzia szidati* miracidia.

As we said, miracidia for a short span should find a mollusc of a certain kind to insinuate into it, otherwise it perishes. Molluscs, thus, are attractants for miracidia. More precisely, the chemotaxis holds as appropriate attractant (a miracidium moves towards a chemical signal proceeding from a mollusc). Other kinds of attractants for miracidia are presented by light (there is a positive phototaxis) and gravitation (negative geotaxis). We will designate all miracidian attractants by $A_1^m, A_2^m, \ldots, A_n^m$.

Miracidian repellents have not been detected still, i.e. $\{R_1^m, R_2^m, \ldots, R_n^m\} = \emptyset$.

In a body of mollusc, a miracidium undergoes metamorphosis and it is transformed into a mother's sporocyst in which a daughter sporocyst educes. In the latter then cercariae start to be formed. This state can be called the *miracidian diffusion*. We will designate these diffusions by $C_1^m, C_2^m, \ldots, C_n^m$.

Now we can construct a version of *Physarum* spatial logic for simulating the behavior of local groups of miracidia. The processes have the following syntax which is defined in the way of *Physarum* logic:

$$P, Q := Nil \mid \gamma.P \mid A^m(p).P \mid \overline{A^m(p)}.P \mid C^m(p).P \mid (P|Q) \mid P \setminus Q \mid P \& Q \mid P + Q \mid a$$

For the simulation, we need, therefore, to have two sets of actions T and T^m, where T contains actions of *Physarum* plasmodium, T^m includes actions of local group of miracidia. These sets must have the same number of members (the same cardinality), namely we must have the same number of active zones (growing

pseudopodia and active groups of miracidia), the same number of attractants to reach diffusions (motions of protoplasmic tubes towards food and miracidian motions towards chemical signals of eventual hosts to transform into a maternal sporocyst). For instance, if we have five molluscs in one experimental dish with water and a suspension of miracidia, then we can try to simulate the miracidian processes by *Physarum* spatial logic through stimuli of five nutrient sources with similar localizations as that of molluscs.

3.2 Cercariae (Genus Trichobilharzia)

3.2.1 Morphology of Cercariae

Cercariae are free-swimming larvae of pubertal generation parasitizing vertebrate animals. They are capable to insinuate into the skin of human being who is for them a casual host, invoking at him/her an allergic reaction, the so-called *cercarial dermatitis*. For the first time this term was used by V.V. Cort [7]. This author first correlated this disease with molluscs of certain kinds, and then with cercariae (free-swimming larvae of pubertal maritas) excreted by them. In the human body, these parasites cannot reach a full growth, dying soon after the penetration [18].

Any identification of specific difference of cercariae of the genus *Trichobilharzia* on morphological characters is extremely complicated because of the identity of their constitution and close dimensional characteristics [35], as well as the similarity of their behavior. They belong to the bunch of Furcocercariae, their posterior tail part consists of two branches (furcae), and the length of furca is approximately a half of length of tail stem. Even in a small zoom its pigmented eye-spots are well visible. The cercarial body is translucent. During motion it is strongly reduced, receiving various forms (fig. 2).

Trichobilharzia szidati has the length about 1.0mm and it is easily distinguished by human eye. Its body is oblong, a bit more thickened in the middle. The tail is without appreciable thickening, its length always is more than the length of body (fig. 3). The two pigmented eye-spots are located dorsally (their distance from the frontal end is about 2/5 of the body length). The distance of ventral sucker (acetabulum) is about 2/3 of the body length from the frontal side. The frontal part of body is occupied by a penetration organ which represents the transformed oral sucker. This pear-shaped formation penetrated by oral opening and penetration gland ducts.

All European species of *Trichobilharzia* possess the five pairs of penetration glands. The two pairs are presented by circumacetabular glands located round the ventral sucker, and the three other by postacetabular glands located sequentially one after another below. The secreta of penetration glands helps larvae to break a dermal barrier of vertebrates host.

The somatotype of excretory system is called “mesostoma” which is characterized by that the main collective canals do not come forward further than the level of ventral sucker, divided into the frontal and back canals of the second order.

Fig. 2 The appearance of *Trichobilharzia szidati*. A detection place: the lake of Naroch (the Minsk region, the Mjadel destrict). *Lymnaea stagnalis* is an intermediate host

In the resting state, cercariae are attached to a vascular wall or on a water film by means of acetabulum. Active motions are characteristic only by the strong shaking of pot or by the water interfusion. At a weak rotation of pot, it is visible that the cercarial body and its tail follow the water stream, while their acetabulums keep cercariae on the pot wall. Any continuous active motion is not observed.

3.2.2 Behavior of Cercariae

The cercarial behavior of bird schistosomes (family Schistosomatidae) is well studied due to representatives of the genus *Trichobilharzia* [19]. Their behavior is characterized by a specific taxis which is referred to an effective search of necessary final hosts. This taxis developed by evolution of larvae of bird schistosomes allows their looking for specific hosts to be successful, forward their affixion to a surface of host body as well as their incorporation into a host cutaneous covering and their penetration into a circulatory system, where a parasite reaches sexual maturity. Thus, taxis forms an enough large family of attractants for cercariae.

Cercariae of the genus *Trichobilharzia* after leaving a mollusc are actively swimming in the water for an hour. Such an active behavior of larvae after leaving a mollusc provides a cercarial allocation in the water space. Then cercariae pass to a passive behavior. They are attached by a ventral sucker to a surface film of water or to various subjects near a water surface, getting thus a characteristic resting state (fig. 4). The resting state allows cercariae in absence of specific to them stimulants to stop the search for a host and to conserve their energy.

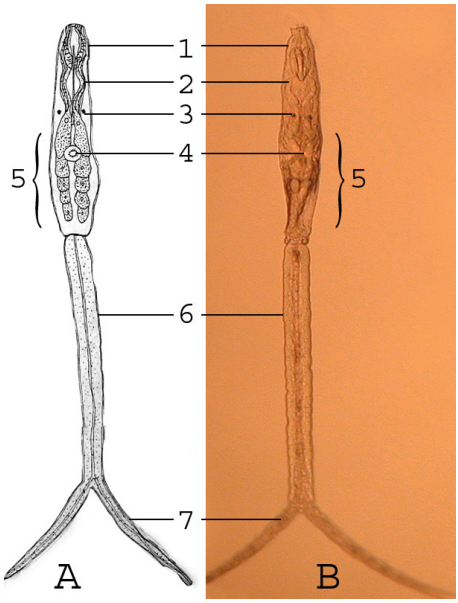


Fig. 3 The constitution of *Trichobilharzia szidati*. A. The schematic structure of cercaria (from [8]), B. The photo of cercaria. 1. Penetration organ, 2. Penetration gland ducts, 3 Pigmented eye-spots, 4. Ventral sucker, 5. Penetration glands (5 pairs), 6. Tail stem, 7. Furcae

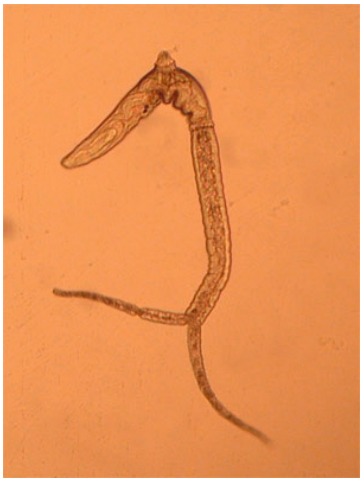


Fig. 4 The resting state of cercaria *Trichobilharzia szidati*. A detection place: the lake of Naroch (the Minsk region, the Mjadel destrict). *Lymnaea stagnalis* is an intermediate host

Free-swimming cercariae need to insinuate into a definitive host during the limited time interval (1–1.5 days at 24°C) since otherwise larvae perish [26].

For a successful search of hosts, larvae of digeneans of *Trichobilharzia* have developed by evolution a behavior facilitating this problem. They possess a positive phototaxis, negative geotaxis, chemotaxis, and also actively react to turbulence of water [11]. It means that for cercariae there are already many other attractants.

The light sensitivity of cercariae of *Trichobilharzia* is very high. As experiments show, cercariae always move towards a light source, and then take a resting state on the lighted side of capacity in which they are. The given taxis, and also negative geotaxis allow cercariae to be kept in the nature at a surface of water in expectation of suitable hosts.

Cercariae actively react to changes in intensity of illumination (shadings) and to turbulence of water [9]. These external factors, corresponding to possible appearances of definitive hosts in water, stimulate the cercarial transition from a resting state into actions that enlarge their chances to meet hosts.

Cercariae possess a chemotaxis in relation to specific hosts. On the body surfaces, larvae of the genus *Trichobilharzia* have chemoreceptors which receive appropriate chemical signals proceeding from a skin of potential host. The similarity of compound of fatty acids of bird and human skin leads to that cercariae equally react to the bird and human appearance in water: they move in their direction, and then they are attached to the skin and begin penetration into it [13]. So, the chemotaxis from a skin of potential hosts (surface lipids of skin of human being and swimming bird), the positive phototaxis, the negative geotaxis and the water turbulence present cercarial attractants of different degree of intensity and power. We will designate these attractants by $A_1^c, A_2^c, \ldots, A_n^c$.

In experimental researches it has been shown that any attachment of cercariae of *Trichobilharzia* to the skin is stimulated by cholesterol and ceramides, and incorporation into skin by linoleic and linolenic acids, all these materials are present in the skin of both bird and human being [12, 24]. Thereby surface lipids of human skin invoke higher frequency of cercarial incorporations into the skin, than surface lipids of birds [13]. One more reason that cercariae of *Trichobilharzia* successfully insinuate into the human skin is the fact of that the skin of duck foots has thicker keratinized surface which, possibly, is more difficult for overcoming, than that of human being [12].

On the basis of experiments, the rate of penetration of larvae of schistosomes *Trichobilharzia szidati* into the human skin [12] has been defined. The larva begins the incorporation into the human skin approximately in 8 seconds (the range from 0 to 80 seconds) after first contacts. The process of full penetration into the skin takes about 4 minutes (the range from 83 seconds to 13 minutes 37 seconds). The given numerals testify that it is enough if the person has even a short-term contact to water where there are cercariae of bird schistosomes to give them possibility to incorporate into the skin.

In cases of human kids, cercarial larvae can penetrate into the skin and be brought by the venous blood to the lungs, invoking there hemorrhages and inflammation. If cercariae are lucky to insinuate into blood and then to lungs, the disease can get harder by the pulmonary syndrome from small cough to symptoms of bronchial obstruction [6].

Thus, there are many attractants for cercariae. At the same time, repellents for them have not been found yet. For example, our experience shows that cercarial motions towards a smaller concentration of material which invokes a destruction of larvae are not observed at all. The experience principle consists in that in a small cavity with the length of 10cm, the width and depth of 0.5cm there is water with a suspension of cercariae. Then a thin essential oil is added in one of the side of this small cavity. Cercariae, which are nearby, quickly perish, although other cercariae do not move aside where the reacting material is absent. Cercariae simply freely float and as soon as they appear in that part, where there is the reacting material, they perish. Thus, $\{R_1^c, R_2^c, \ldots, R_n^c\} = \emptyset$.

In final hosts, cercariae reach diffusion states. We will designate these diffusions by $C_1^c, C_2^c, \ldots, C_n^c$.

The behavior of local groups of cercariae can be simulated within a version of *Physarum* spatial logic, where the processes have the following syntax:

$$P, Q := Nil \mid \gamma.P \mid A^c(p).P \mid \overline{A^c(p)}.P \mid C^c(p).P \mid (P|Q) \mid P \setminus Q \mid P\&Q \mid P+Q \mid a$$

The sets of actions T and T^c, where T consists of actions of *Physarum* plasmodium, T^c contains actions of local group of cercariae, must have the same number of members. For example, if we have three human beings in one lake with cercariae, then we can simulate the cercarial processes by *Physarum* spatial logic where three nutrient sources with similar localizations as that for human beings act as stimuli. Hence, the behavior of local groups of cercariae is another biological implementation of Kolmogorov-Uspensky machines. It can build planar graphs as well.

4 Arithmetic Operations in *Physarum* Spatial Logic and in Schistosomatidae Behavioral Logic

We know that within process calculus we can convert expressions from λ-calculus. In particular, it means that we can consider arithmetic operations as processes. *Physarum* spatial logic as well as its modification in the form of behavioral logic for local groups of miracidia (cercariae) is a biologized version of process calculus. Therefore we can convert arithmetic operations into processes of either *Physarum polycephalum* spatial logic or Schistosomatidae behavioral logic.

Indeed, growing pseudopodia may represent a natural number n by the following parametric process:

$$\underline{n}(x,z) := \underbrace{\overline{x}.\overline{x}.\cdots.\overline{x}}_{n}.\overline{z}.Nil$$

The process $\underline{n}(x,z)$ proceeds n times on an output port called the successor channel $\overline{x} \in \{A_1, A_2, \ldots\} \cup \{R_1, R_2, \ldots\}$ (e.g. it is the same output of attractant) and once on the zero output port $\overline{z} \in \{A_1, A_2, \ldots\} \cup \{R_1, R_2, \ldots\}$ before becoming inactive Nil. Recall that it is a "Church-like" encoding of numerals used first in λ-calculus. Notice that in case of miracidia or cercariae $\overline{x} \in \{A_1, A_2, \ldots\}$ and $\overline{z} \in \{A_1, A_2, \ldots\}$.

An addition process takes two natural numbers i and j represented using the channels $x[i], z[i]$ and $x[j], z[j]$ and returns their sum as a natural number represented using channels $x[i+j], z[i+j]$:

$Add(x[i], z[i], x[j], z[j], x[i+j], z[i+j]) := (x[i].\overline{x}[i+j].\ Add(x[i],\ z[i],\ x[j],\ z[j],\ x[i+j],\ z[i+j])) + z[i].Copy(x[j], z[j],\ x[i+j],\ z[i+j]))$.

A multiplication process takes two natural numbers i and j represented using the channels $x[i], z[i]$ and $x[j], z[j]$ and returns their multiplication as a natural number represented using channels $x[\underbrace{i+\ldots+i}_{j}]$, $z[\underbrace{i+\ldots+i}_{j}]$:

$Mult(x[i], z[i],\ x[j], z[j],\ x[i*j],\ z[i*j]) := Add(x[i], z[i],\ x[j], z[j],\ x[i+\ldots+i],\ z[i+\ldots+i])$.

The *Copy* process replicates the signal pattern on channels x and y on to channels u and v. It is defined as follows:

$Copy(x, y, u, v) := (x.\overline{u}.Copy(x, y, u, v) + y.\overline{v}.Nil)$

As we see, within *Physarum* spatial logic and its poorer version in the form of Schistosomatidae behavioral logic, we can consider some processes as arithmetic operations. Also, we can combine several arithmetic operations within one process. Let us regard the following expression:

$$(10+20)*(30+40)$$

An appropriate process is as follows:

$Mult(Add(x[10], z[10], x[20], z[20], x[10+20], z[10+20]), z[10+20], Add(x[30],\ z[30],\ x[40], z[40],\ x[30+40], z[30+40]), z[30+40], Add(x[30], z[30], x[70],\ z[70],\ x[2100], z[2100]), z[2100])$.

5 *Physarum polycephalum* Illocutionary Logic and Schistosomatidae Illocutionary Logic

Involving *Physarum* process calculus as a programming language in analyzing intelligent processes allows us to formalize many kinds of human interactions within *Physarum* automata. For instance, it is possible to consider simpler versions of *illocutionary logic* that was developed for explicating the logical nature of human speech acts and for checking the illocutionary Turing test on the medium of *Physarum polycephalum* plasmodia or miracidia or cercariae collectives. This logic studies *illocutionary propositions* which express our emotional and cognitive valuations to commit interactions. Let us show that indeed, in *Physarum* process calculus, we can logically formulate some simple human illocutionary propositions.

Suppose, Ψ is any proposition that is built up by a superposition of standard propositional logical connectives ($\wedge$, $\vee$, $\neg$, $\Rightarrow$) in the conventional way. Let V be a valuation of each propositional variable p such that $V(p) \subseteq \mathscr{L}$. We mean that $V(p)$ consists of possible worlds, where p is true. Now we can define whether a proposition Ψ is true in the event x of $\mathscr{L}$. If it is true, it is denoted by $x \models \Psi$.

$x \models p$ if p belongs to $V(p)$;

not $x \models false$, where *false* is any contradiction;

$x \models (\Psi \Rightarrow \Phi)$ iff not $x \models \Psi$ or $x \models \Phi$;

$x \models \neg\Psi$ iff $x \models (\Psi \Rightarrow false)$.

Within *Physarum* spatial logic, we can define the following three illocutionary propositions:

$\bigcirc_{eat}\Psi$:= '*I would like to eat* Ψ'

$\bigcirc_{fear}\Psi$:= '*I fear* Ψ'

$\bigcirc_{satisfy}\Psi$:= '*I am satisfied by* Ψ'

These propositions have the following semantics:

$y \models \bigcirc_{eat}\Psi$ if for any process P containing $\overline{A(x)}$, we have that if $x \models \Psi$, then P contains a transition $x \overset{\overline{A(x)}}{\rightarrow} y$

$y \models \bigcirc_{fear}\Psi$ if for any process P containing $\overline{R(x)}$, we have that if $x \models \Psi$, then P contains a transition $x \overset{\overline{R(x)}}{\rightarrow} y$.

$x \models \bigcirc_{satisfy}\Psi$ if for any process P containing $C(x)$, we have that if P contains a transition $x \overset{C(x)}{\rightarrow} y$, then $y \models \Psi$.

$\Diamond_{eat}\Psi := \neg(\bigcirc_{eat}(\neg\Psi))$.

$\Diamond_{fear}\Psi := \neg(\bigcirc_{fear}(\neg\Psi))$.

$\Diamond_{satisfy}\Psi := \neg(\bigcirc_{satisfy}(\neg\Psi))$.

As we see, we can consider the plasmodium activity as a verification of three basic illocutionary propositions: $\bigcirc_{eat}\Psi$, $\bigcirc_{fear}\Psi$, and $\bigcirc_{satisfy}\Psi$ (in case of miracidia or cercariae we have just two basic illocutionary propositions: $\bigcirc_{eat}\Psi$ and $\bigcirc_{satisfy}\Psi$).

Meanwhile, usually two or more different localizations of *Physarum polycephalum* do not compete with each other and reach the Nash equilibrium soon. For instance, their illocutionary propositions '*I would like to eat*' may be regarded as their joint proposition. The illocutionary Turing test means that we have a verification of illocutionary acts $\bigcirc_{eat}\Psi$, $\bigcirc_{fear}\Psi$, and $\bigcirc_{satisfy}\Psi$ for plasmodia of *Physarum polycephalum* as well as a verification of illocutionary acts $\bigcirc_{eat}\Psi$ and $\bigcirc_{satisfy}\Psi$ for miracidia and cercariae collectives.

6 Conclusion

We have shown that many biologized versions of process calculus are possible: *Physarum polycephalum* spatial logic, Schistosomatidae behavioral logic, etc., and appropriate modal logics for verifying speech acts: *Physarum polycephalum* illocutionary logic, Schistosomatidae illocutionary logic, etc. One of its basic versions, *Physarum* spatial logic, can be used for constructing slime mould based computer. This logic is richer than Schistosomatidae behavioral logic and may be involved for simulations of the latter as well as *Physarum polycephalum* illocutionary logic is reacher than Schistosomatidae illocutionary logic. The fact that we can formalize biological behaviors as kind of logic confirms that biological processes can be considered as forms of concurrent computations.

Acknowledgements. This research is being fulfilled by the support of FP7-ICT-2011-8 and UMO-2012/07/B/HS1/00263.

References

1. Adamatzky, A., De Lacy Costello, B., Asai, T.: Reaction-Diffusion Computers. Elsevier, Amsterdam (2005)
2. Adamatzky, A.: Physarum Machines: Computers from Slime Mould (World Scientific Series on Nonlinear Science, Series A). World Scientific Publishing Company (2010)
3. Adamatzky, A.: Physarum machine: implementation of a Kolmogorov-Uspensky machine on a biological substrate. Parallel Processing Letters 17(04), 455–467 (2007)
4. Adamatzky, A., Erokhin, V., Grube, M., Schubert, T., Schumann, A.: Physarum Chip Project: Growing Computers From Slime Mould. International Journal of Unconventional Computing 8(4), 319–323 (2012)
5. Adamatzky, A., Akl, S., Alonso-Sanz, R., van Dessel, W., Ibrahim, Z., Ilachinski, A., Jones, J., Kayem, A., Martinez, G.J., de Oliveira, P., Prokopenko, M., Schubert, T., Sloot, P., Strano, E., Yang, X.: Are motorways rational from slime mould's point of view? Int. J. of Parallel, Emergent and Distributed Systems 28(3), 230–248 (2012)
6. Bear, S.A., Voronin, M.V.: Cercariae in Urbanized Ecosystems. Nauka, Moskow (2007) (in Russian)
7. Cort, W.W.: Schistosome dermatitis in the United States (Michigan). Journal of the American Medical Association 90, 1027–1029 (1928)
8. Ginetsinskaya, T.A.: Trematodes, their Life Cycles, Biology and Evolution. Nauka, Leningrad (1968)

9. Feiler, W., Haas, W.: Trichobilharzia ocellata: chemical stimuli of duck skin for cercarial attachment. Parasitology 96, 507–517 (1988)
10. Haas, W.: Host finding mechanisms – a physiological effect. In: Mehlhorn, Y. (ed.) Biology, Structure, Function: Encyclopedic Reference of Parasitology, 2nd edn., pp. 382–383 (1988)
11. Haas, W.: Physiological analysis of cercarial behavior. Journal of Parasitology 78, 243–255 (1992)
12. Haas, W., Haeberlein, S.: Penetration of cercariae into the living human skin: Schistosoma mansoni vs. Trichobilharzia szidati. Parasitology Research, 1061–1066 (2009)
13. Haas, W., Roemer, A.: Invasion of the vertebrate skin by cercariae of *Trichobilharzia ocellata*: penetration processes and stimulating host signals. Parasitology Research 84(10), 787–795 (1998)
14. Haas, W.: Parasitic worms: strategies of host finding, recognition and invasion. Zoology 106, 349–364 (2003)
15. Haas, W., Haberl, B., Kalbe, M., Körner, M.: Snail-host-finding by Miracidia and Cercariae: Chemical Host Cues. Parasitology Today 11, 468–472 (1995)
16. Haas, W., Haberl, B.: Host recognition by trematode miracidia and cercariae. In: Advances in Trematode Biology, pp. 197–227. CRC Press, Boca Raton (1997)
17. Hertel, J., Holweg, A., Haberl, B., Kalbe, M., Haas, W.: Snail odour-clouds: spreading and contribution to the transmission success of Trichobilharzia ocellata (Trematoda, Digenea) miracidia. Oecologia 147, 173–180 (2006)
18. Horák, P., Kolárová, L.: Bird schistosomes: do they die in mammalian skin? Trends in Parasitology 17(2), 66–69 (2001)
19. Horák, P., Kolárová, L., Adema, C.M.: Biology of the schistosome genus *Trichobilharzia,*. Advances in Parasitology 52, 155–233 (2002)
20. Jones, J.: Towards Programmable Smart Materials: Dynamical Reconfiguration of Emergent Transport Networks. Int. Journal of Unconventional Computing 7(6), 423–447 (2011)
21. Kalbe, M., Haberl, B., Haas, W.: Finding of the snail host by Fasciola hepatica and Trichobilharzia ocellata: compound analysis of Miracidia attracting glycoprotein. Experimental Parasitology 96, 231–242 (2000)
22. Kock, S.: Investigations on intermediate host specificity help to elucidate the taxonomic status of *Trichobilharzia ocellata* (Digenea: Schistosomatidae). Parasitology 123, 67–70 (2001)
23. MacInnis, A.J.: How parasites find their hosts: some thoughts on the inception of host-parasite integration. In: Cennedy, C.R. (ed.) Ecological Aspects of Parasitology, pp. 3–20. North-Holand, Amsterdam (1976)
24. Mikes, L., Zìdková, L., Kasný, M., Dvorák, J., Horák, P.: In vitro stimulation of penetration gland emptying by Trichobilharzia szidati and T. regenti (Schistosomatidae) cercariae. Quantitative collection and partial characterization of the products. Parasitology Research 96(4), 230–241 (2005)
25. Molyneux, D.H.: Control of human parasitic diseases: context and overview. Advances in Parasitology 61, 1–45 (2006)
26. Neuhaus, W.: Biologie und Entwicklund von *Trichobilharzia szidati* n. sp (Trematoda, Schistosomatidae), einem Erreger von Dermatitis beim Menschen. Zeitschrift für Parasitenkunde 15, 203–266 (1952)
27. Saladin, K.S.: Behavioral parasitology and perspectives on miracidial host-finding. Zeitschrift für Parasitenkunde 3, 197–210 (1979)

28. Schumann, A., Adamatzky, A.: Logical Modelling of Physarum Polycephalum. Analele Universitatii de Vest, Timisoara, Seria Matematica – Informatica XLVIII 3, 175–190 (2010)
29. Schumann, A., Adamatzky, A.: Physarum Spatial Logic. New Math. and Nat. Computation 7(3), 483–498 (2011)
30. Searle, J.R., Vanderveken, D.: Foundations of Illocutionary Logic. Cambridge University Press, Cambridge (1984)
31. Smyth, J.D., Halton, D.W.: The Physiology of Trematodes, 2nd edn. Cambridge University Press, Cambridge (1983)
32. Sukhdeo, M.V.K., Mettrick, D.F.: Parasite behaviour: understanding platyhelminth responses. Advances Parasitology 26, 73–144 (1987)
33. Sukhdeo, M.V.K., Sukhdeo, S.C.: Trematode behaviours and the perceptual worlds of parasites. Canadian Journal of Zoology 82, 292–315 (2004)
34. Takahashi, T., Mori, K., Shigeta, Y.: Phototactic, thermotactic and geotactic responses of miracidia of *Schistosoma japonicum*. Japanese Journal of Parasitology 110, 686–691 (1961)
35. Podhorsky, M., Huzova, Z., Mikes, L., Horak, P.: Cercarial dimensions and surface structures as a tool for species determination of Trichobilharzia spp. Acta Parasitologica 50, 343–365 (2009)

Towards Lateral Inhibition and Collective Perception in Unorganised Non-neural Systems

Jeff Dale Jones

Abstract. Lateral Inhibition (LI) phenomena occur in a wide range of neural sensory modalities, but are most famously described in the visual system of humans and other animals. The general mechanism can be summarised as when a stimulated neuron is itself excited and also suppresses the activity of its local neighbours via inhibitory connections. The effect is to generate an increase in contrast between spatial environmental stimuli. Simple organisms, such as the single-celled slime mould *Physarum polycephalum* possess no neural tissue yet, despite this, are known to exhibit complex computational behaviour. Could simple organisms such as slime mould approximate LI without recourse to neural tissue? We describe a model whereby LI can emerge without explicit inhibitory wiring, using only bulk transport effects. We use a multi-agent model of slime mould to reproduce the characteristic edge contrast amplification effects of LI using excitation via attractant based stimuli. We also explore a counterpart behaviour, Lateral Activation (where stimulated regions are inhibited and lateral regions are excited), using simulated exposure to light irradiation. In both cases restoration of baseline activity occurs when the stimuli are removed. In addition to the enhancement of local edge contrast the long-term change in population density distribution corresponds to a collective response to the global brightness of 2D image stimuli, including the scalloped intensity profile of the Chevreul staircase and the perceived difference of two identically bright patches in the Simultaneous Brightness Contrast (SBC) effect. This simple model approximates LI contrast enhancement phenomena and global brightness perception in collective unorganised systems without fixed neural architectures. This may encourage further research into unorganised analogues of neural processes in simple organisms and suggests novel mechanisms to generate collective perception of contrast and brightness in distributed computing and robotic devices.

Jeff Dale Jones
Centre for Unconventional Computing,
University of the West of England, Bristol, UK
e-mail: jeff.jones@uwe.ac.uk

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*,
Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_6

1 Introduction

Living organisms perceive their environment with a wide variety of special sensory modalities. Enhancing the contrast in the stream of information from these senses allows organisms to discriminate between small changes in signal level, potentially enhancing survivability. Lateral Inhibition (LI) is a neural mechanism which enhances the activity of neurons directly exposed to excitatory stimuli whilst suppressing the activity of their near neighbours (see Fig. 1 for a schematic illustration). LI phenomena have been described in auditory [9], somatosensory [22] and olfactory senses [29], but are most famously described in the visual systems of a wide range of animals, including humans [8, 15, 14].

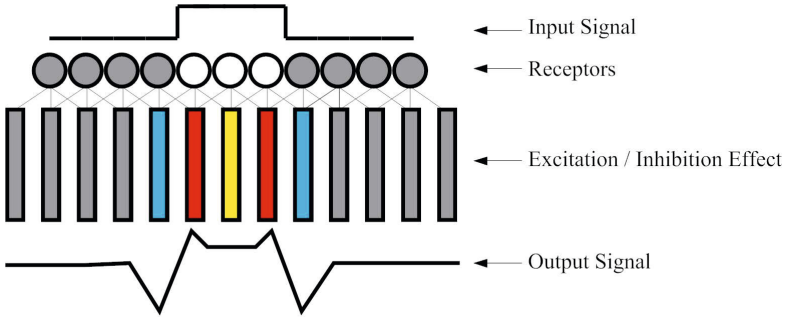


Fig. 1 Schematic illustration of Lateral Inhibition response. Original stimulus (top) activates neurons (circled) whose local inhibitory wiring results in excitation and inhibition effects in neighbouring neurons (bars), resulting in the enhanced output signal (bottom)

2 Non-neural Computation in Slime Mould

LI phenomena, and the mechanisms which generate them, result in an effective and efficient means of enhancing environmental perception in organisms containing nervous systems ranging from the most complex to the most primitive. But can such mechanisms occur in organisms which do not possess any nervous system? The true slime mould *Physarum polycephalum* is a single-celled amoeboid organism with a very complex life cycle. The plasmodium stage, where a giant syncytium formed by repeated nuclear division is encompassed within a single membrane, has been shown to exhibit a complex range of biological and computational behaviours. The plasmodium of *P. polycephalum* is comprised of a complex gel/sol transport network which is continuously remodelled in response to its environment. The organism behaves as a distributed computing *material*, capable of responding to a wide range of spatially represented stimuli. The ectoplasmic gel phase is composed of a sponge-like matrix of contractile actin and myosin fibres through which the endoplasmic sol

flows. Local oscillations in the thickness of the plasmodium spontaneously appear with approximately 2 minutes duration [24]. The spatial and temporal organisation of the oscillations has been shown to be extremely complex [25] and affects the internal movement of sol through the network by local assembly and disassembly of the actin-myosin structures. The protoplasm moves backwards and forwards within the plasmodium in a characteristic manner known as shuttle-streaming.

The plasmodium is able to sense local concentration gradients and the presence of nutrient gradients appears to alter the structure of external membrane areas. The softening of the outer membrane causes a flux of protoplasm towards the general direction of the gradient in response to internal pressure changes caused by the local thickness oscillations. The strong coupling between membrane contraction and streaming movement is caused by the incompressibility of the fluid requiring a constant volume - the weakening of the membrane provides an outlet for the pressure. When the plasmodium has located and engulfed nearby food sources the material behind the solid active growth front begins to form a network of protoplasmic veins which become coarser in distal regions. These veins connect nutrient sources in networks that have been shown to be efficient in terms of their overall distance and resilience to random damage [16] (see Fig. 2 for an example transport network connecting oat flake nutrient sources).

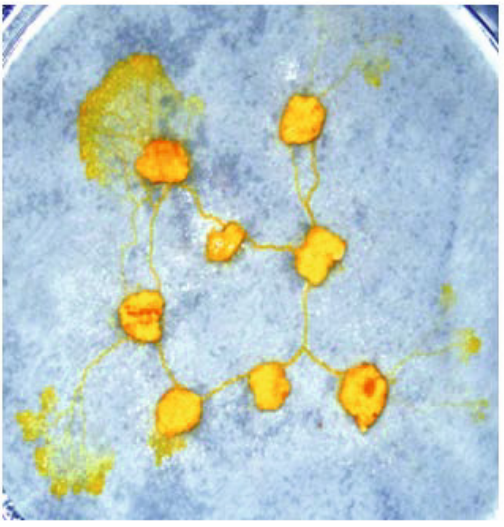


Fig. 2 Plasmodium of slime mould *Physarum polycephalum* inoculated on damp filter paper forms a transport network connecting oat flake nutrients. Note the fan-shaped solid active growth front of the plasmodium at the top-left of the image, with the network forming and coarsening behind the active front

The transport network distributes protoplasm (including microscopic nutrients) throughout the organism. The relative simplicity of the cell and the distributed nature of its control system make *P. polycephalum* a suitable subject for research into distributed computation substrates. In recent years there have been a large number of studies investigating its computational abilities, prompted by Nakagaki et al. who reported the ability of *P. polycephalum* to solve path planning problems [17]. Subsequent research confirmed and broadened the range of abilities to spatial representations of various graph problems [16, 23, 2], combinatorial optimisation problems [5], construction of logic gates [27] and logical machines [1], [4], and as a means to achieve distributed robotic control [28], robotic manipulation [3] and robotic amoeboid movement [10], [13].

In this paper we take inspiration from very simple organisms, such as slime mould, without neural tissue which nevertheless exhibit very complex behaviours. Although *Physarum* does not possess neural tissue, it displays behaviour which is analagous to neural spiking and synaptic learning [7], [19]. *Physarum* has been used to study potential low-level mechanisms of perception in the Kanizsa illusory contour phenomena [26] and other simple collective organisms have been used to study geometric illusions, such as the Muller-Lyer effect [21]. In this paper we initially start from a simpler level, describe collective mechanisms by which sensory contrast enhancement phenomena analogous to Lateral Inhibition phenomena can emerge in disorganised non-neural systems.

We use a multi-agent particle based model of slime mould to demonstrate and elucidate the low-level behaviours which generate these collective phenomena. In Section 3 we describe the multi-agent model of slime mould, its previous use as a material computing substrate, and parameters governing its use for the experiments in this chapter. In Section 4 we examine the emergence of LI phenomena in the model in response to presentation with attractant stimuli. The opposite response — Lateral Activation — is described in response to presentation with adverse stimuli (simulated light irradiation) in Section 5. In Section 6 we build on these simple mechanisms by showing how a *global* collective representation of an environment (in this case an approximation of overall brightness) can be generated by bulk drift of population density in response to greyscale spatial patterns presented as attractant stimuli. We conclude in Section 7 by summarising the results, the main contributions of this chapter, and examining potential applications of these collective phenomena for computing and robotics applications.

3 Multi-agent Model of Slime Mould

We used a multi-agent approach to generate the *Physarum*-like behaviour. This approach was chosen specifically specifically because we wanted to reproduce the generation of complex behaviour in slime mould using the same limitations that slime mould has, i.e. using very simple component parts and interactions, and no special or critical component parts to generate the emergent behaviour. Although other

modelling approaches, notably cellular automata, also share these properties, the direct mobile behaviour of the agent particles renders it more suitable to reproduce the flux within the *Physarum* plasmodium. The multi-agent particle model [12] uses a population of coupled mobile particles with very simple behaviours, residing within a 2D diffusive lattice. The lattice stores particle positions and the concentration of a local diffusive factor referred to generically as chemoattractant. Particles deposit this chemoattractant factor when they move and also sense the local concentration of the chemoattractant during the sensory stage of the particle algorithm. Collective particle positions represent the global pattern of the material. The model runs within a multi-agent framework running on a Windows PC system. Performance is thus influenced by the speed of the PC running the framework. The particles act independently and iteration of the particle population is performed randomly to avoid any artifacts from sequential ordering.

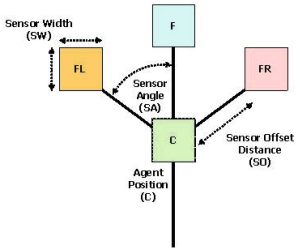
3.1 Generation of Collective Cohesion Phenomena

The behaviour of individual particles occurs in two distinct stages, the sensory stage and the motor stage. In the sensory stage, the particles sample their local environment using three forward biased sensors whose angle from the forwards position (the sensor angle parameter, SA), and distance (sensor offset, SO) may be parametrically adjusted (Fig. 3a). The offset sensors generate local indirect coupling of sensory inputs and movement to generate the cohesion of the material. The SO distance is measured in pixels and a minimum distance of 3 pixels is required for strong local coupling to occur. It was shown in [11] that large SO values result in regular self-organised domains (‘vacancy islands’) of small vacant regions in the material. This would not be desirable in these experiments as we gauge the response of the material by measuring population distribution density (which would be affected by the presence of these domains), so we randomly selected the SO parameter for each particle, at each scheduler step, from the range of 1 – 20 pixels. This maintained the cohesion of the material whilst avoiding the formation of these vacancy domains. During the sensory stage each particle changes its orientation to rotate (via the parameter rotation angle, RA) towards the strongest local source of chemoattractant (Fig. 3b). Variations in both SA and RA parameters have been shown to generate a wide range of reaction-diffusion patterns [11] and for these experiments we used SA 60 and RA 60. After the sensory stage, each particle executes the motor stage and attempts to move forwards in its current orientation (an angle from 0–360 degrees) by a single pixel forwards. Each lattice site may only store a single particle and particles deposit chemoattractant into the lattice (5 units per step) only in the event of a successful forwards movement. If the next chosen site is already occupied by another particle move is abandoned and the particle selects a new randomly chosen direction.

3.2 Representation of Spatial Stimuli

The presentation of simple uniform attractant stimuli to the virtual plasmodium was achieved by incrementing attractant values by 1.275 units at stimulated regions every scheduler step. The presentation of Adverse stimuli (simulated light irradiation) to the model was achieved by reducing the sensitivity of the particle sensors in illuminated regions by 80% and reducing particle chemoattractant deposition in the same regions by 80%.

Representation of more complex stimuli with difference brightness levels (Section 6) was achieved by incrementing attractant values by the corresponding stimuli image pixel brightness at each location on the lattice and scaling this value downwards by multiplying by 0.01. The reduction in attractant concentration by scaling reduces the attractant stimuli concentration towards the baseline flux generated by the particle movement and maintains the integrity of the virtual plasmodium (stronger stimuli would cause the material to tear). Diffusion within in the lattice was implemented at each scheduler step and at every site in the lattice via a simple mean filter of kernel size 5×5. Damping of the diffusion distance, which limits the distance of chemoattractant gradient diffusion, was achieved by multiplying the mean kernel value by 0.95 per scheduler step.



(a)

```
[Sensory stage]

- Sample chemoattractant map values
- if (F > FL) && (F > FR)
       - Continue facing same direction
- Else if (F < FL) && (F < FR)
       Rotate by RA towards larger of FL and FR
- Else if (FL < FR)
       Rotate right by RA
- Else if (FR < FL)
       Rotate left by RA
- Else
       Continue facing same direction
```

(b)

Fig. 3 Architecture of a single particle of the virtual material and its sensory algorithm: (a) Morphology showing agent position 'C' and offset sensor positions (FL, F, FR), (b) Algorithm for particle sensory stage

3.3 Translation of Neuronal Responses to Non-neural Mechanisms

To explore mechanisms corresponding to LI in non-neural systems we must translate the relevant neuroscience terminology into terms which can be represented in the model system. Input stimuli (such as light to the human visual system) can be represented in the model by spatial projection of simulated chemoattractants. These

stimuli attract the particles comprising the virtual plasmodium. We may say that the projection of chemoattractants results in an excitatory response. Conversely we can generate an inhibitory response by the spatial projection of repellents or other adverse stimuli. Slime mould is known to avoid illumination with certain wavelengths of visible light and we can use this feature to generate an inhibitory response in the model.

The basic responses of excitation (by chemoattraction) and inhibition (by repulsion or irradiation) are short term approximations of the neural response. To approximate long-term neural responses (changes in spatio-temporal patterns of neural activity) we must utilise the changing spatial density distribution of the particles comprising the model material, in response to the projected stimuli, i.e. changes in the population density correspond to regions of increased / decreased neural activity. The model must also respond to the withdrawal of stimuli. In neural systems this would result in a reversion to baseline activity. In the model this must be represented by the restoration of uniform population density on withdrawal of the stimuli.

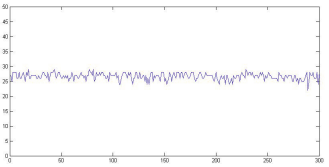
4 Lateral Inhibition Phenomena Using Attractant Stimuli

We initialised the virtual plasmodium comprising 8000 particles within a 300×100 pixel tube-like horizontal arena bordered by inhabitable areas on the top and bottom and open ended left and right edges. Periodic boundary conditions were enforced. We measured population density across the arena by counting the number of particles in the Y-axis for each X-axis position. We recorded population density every 10 scheduler steps. After initialisation the population, constrained by the architecture of the arena, formed a single tube with relatively uniform population density (Fig. 4,a).

An attractant stimulus was presented to the virtual plasmodium after 500 scheduler steps by projecting chemoattractant into the middle-third habitable section of the arena (white region in Fig. 4,c). The attractant stimulus caused increased flux of particles into the stimulus area, an increase in population density in this area, and a corresponding decrease in density outside the stimulus region (Fig. 4,e). Upon removal of the stimulus after 4000 steps the population was no longer attracted to the central region and the tube adapted its shape in response to the uniform chemoattractant profile (Fig. 4,g). The population density eventually returned to uniform density across the arena (Fig. 4,i).

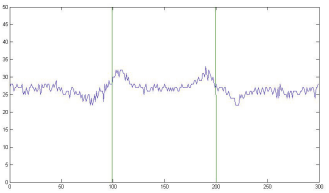
A space-time plot of the population density indicates how the changes in population density are initiated at the stimulus boundaries and propagate outwards from these regions (Fig. 5). Regions inside the attractant stimulus area correspond to excitation areas and regions outside correspond to inhibited activity. The presentation of attractant stimuli, and the response of the particle population, causes an increase in signal contrast (measured in terms of population density) between stimulated and non-stimulated areas (Fig. 6).

(a) t=500



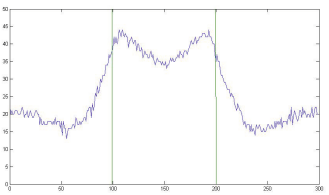
(b) density

(c) t=600



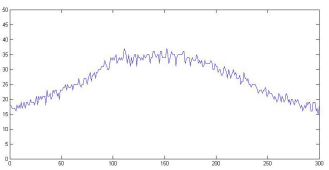
(d) density

(e) t=4000



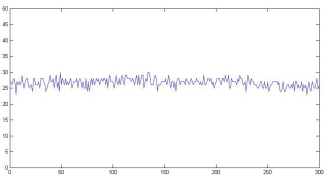
(f) density

(g) t=8000



(h) density

(i) t=20000



(j) density

Fig. 4 Response of virtual plasmodium flux to attractant stimulus: (a) population initialised within horizontal arena forms single tube, (c) presentation of attractant stimulus bar (light area) results in flux towards stimulus area, (e) population density is increased at stimulus region and reduced at unstimulated region, (g) removal of stimulus results in adaptation to uniform attractant profile, (i) uniform density is restored, (b,d,f,h,j) cross-section plots of population density across the arena

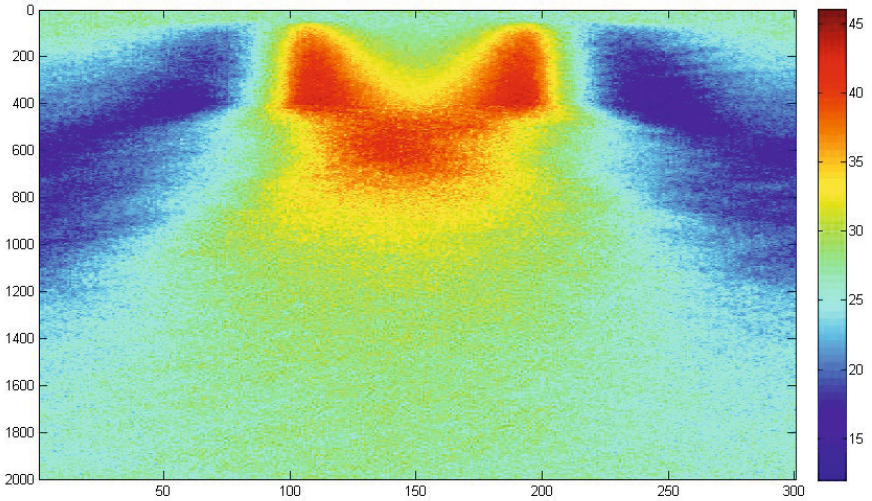


Fig. 5 Space-time plot of population density flux under attractant stimuli conditions, time proceeds downwards. Note that changes in density are initiated at borders of the stimulus boundary

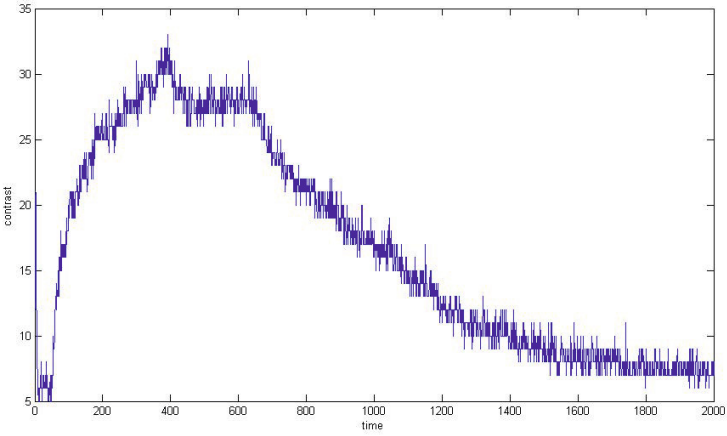
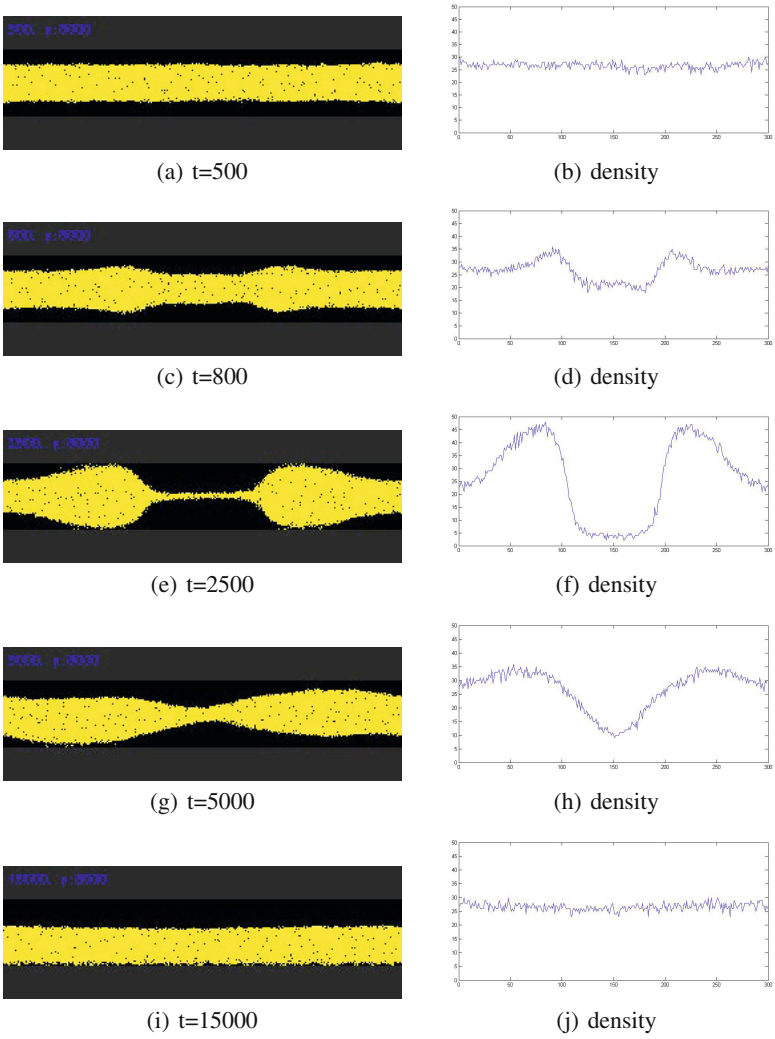


Fig. 6 Presentation of attractant stimuli and its effect on population density, interpreted as signal contrast enhancement. Stimulus presented after 50 samples (t=500), stimulus removed after 400 samples (t=4000) causing gradual reversion to baseline activity



(a) t=500 (b) density

(c) t=800 (d) density

(e) t=2500 (f) density

(g) t=5000 (h) density

(i) t=15000 (j) density

Fig. 7 Response of virtual plasmodium flux to simulated light irradiation: (a) population initialised within horizontal arena forms single tube, (c) presentation of simulated light irradiation (centre, not shown) results in flux away from irradiated area, (e) population density is decreased at irradiated region and increased at unexposed region, (g) removal of adverse stimulus results in increased flux to inner region, (i) uniform density is restored, (b,d,f,h,j) cross-section plots of population density across the arena

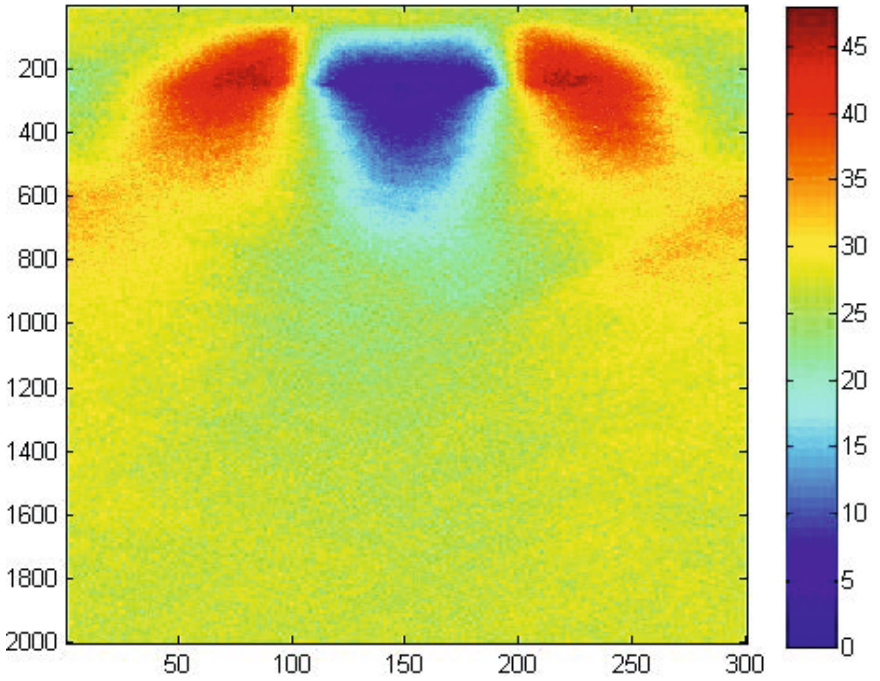


Fig. 8 Space-time plot of population density flux under adverse stimuli condition, time proceeds downwards. Note that changes in density are initiated at borders of stimulus boundary

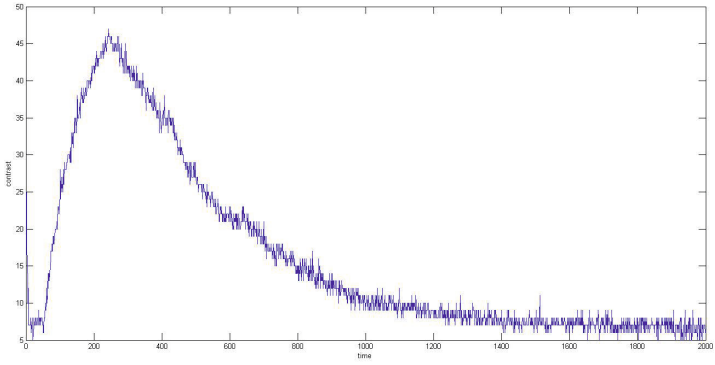


Fig. 9 Presentation of adverse stimuli (simulated light irradiation) and its effect on population density, interpreted as signal contrast enhancement. Adverse stimulus presented after 50 samples (t=500), stimulus removed after 250 samples (t=2500) causing gradual reversion to baseline activity

5 Lateral Activation Phenomena Using Adverse Stimuli

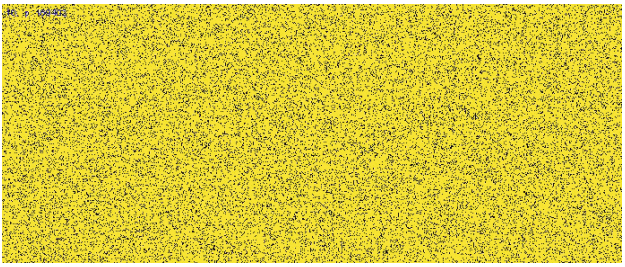
To examine the collective response to Adverse Stimuli (simulated exposure to illumination, which the *Physarum* plasmodium avoids), we used the same arena with the virtual plasmodium inoculated as a horizontal strip with approximately uniform density (Fig. 7, a and density profile in b). The stimulus pattern was again in the central third of the arena, but the attractant region was replaced with a region of simulated illumination. Particles at the border of exposed areas preferentially moved to unexposed regions and the local coupling of particles resulted in collective flux away from the stimulus area (Fig. 7, c and e). The resulting density profile (Fig. 7, d and f) demonstrates the inhibition effect within the illuminated region whilst the unexposed neighbouring regions show an increase in population density. When the adverse stimulus was removed from the central region the population density renormalised to a uniform level within 15000 steps (Fig. 7, g and i and corresponding density profiles in h and j respectively). The space-time plot of changing population density shows the inhibition effect in the central region and the lateral propagation of increased density (Fig. 8). Areas inside the adverse stimulus correspond to an inhibition response and areas outside correspond to excited activity. As in the attractant stimuli case, the presentation of adverse stimuli increases the signal contrast (population density) between illuminated and non-illuminated regions (Fig. 9).

6 From Low-Level Mechanisms to Unorganised Collective Perception

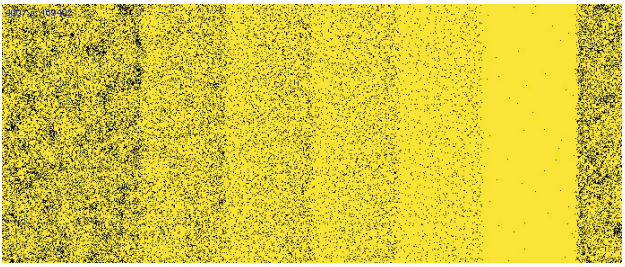
The changes in population density over time in response to patterns of attractant and adverse stimuli correspond to unorganised Lateral Inhibition and Lateral Activation mechanisms respectively. How do these mechanisms respond to more complex arrangements of stimuli? We examine the response of the virtual plasmodium to attractant stimuli in the pattern of the Chevreul staircase illusion (Fig. 10a). The Chevreul staircase is a sequence of identical width uniform vertical bars. Each bar is lighter in intensity then its leftmost neighbour (see cross-section intensity plot in Fig. 10e). Although each bar is uniform in intensity the image is typically perceived as having a scalloped profile across each bar, i.e. the left side of each bar (when adjacent to a darker bar) is perceived as being lighter, and the right side of each bar (when adjacent to a lighter bar) is perceived as being darker. The mechanisms underlying this illusory percept are considered to be mediated by activity at the retinal and cortical levels [18], [20].

We initialised the virtual plasmodium (comprising 169,402 particles) on an arena patterned with the Chevreul staircase image (692×288 pixels). Each vertical bar corresponded to increasing concentration stimuli. Periodic boundary conditions were used and two bordering bars offering no stimuli were placed at the left and right of the image. The initially uniform distribution of particles (Fig. 10b) was affected

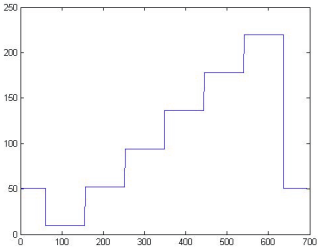
(a) Chevreul Staircase



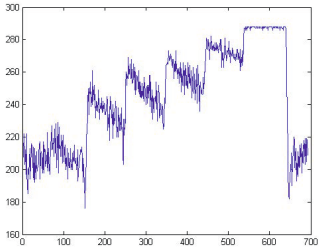
(b) t=10



(c) t=400



(d) original section



(e) population density t=400

Fig. 10 Collective representation of the Chevreul staircase illusion: (a) Original greyscale Image of uniform bars of increasing lightness (surrounded by left and right borders), (b) initial uniform population distribution of particle population, (c) population distribution after 4000 scheduler steps showing increased population density at lighter (greater attractant concentration) regions, (d) cross-section profile of original image stimulus, (e) cross-section of population density at t=400 showing scalloped borders and increasing contrast

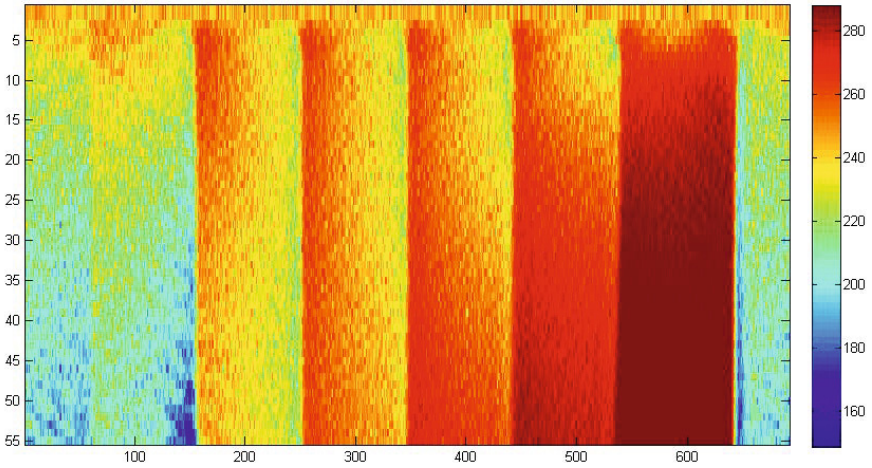


Fig. 11 Space-time plot of population density flux under Chevreul illusory stimulus, time proceeds downwards

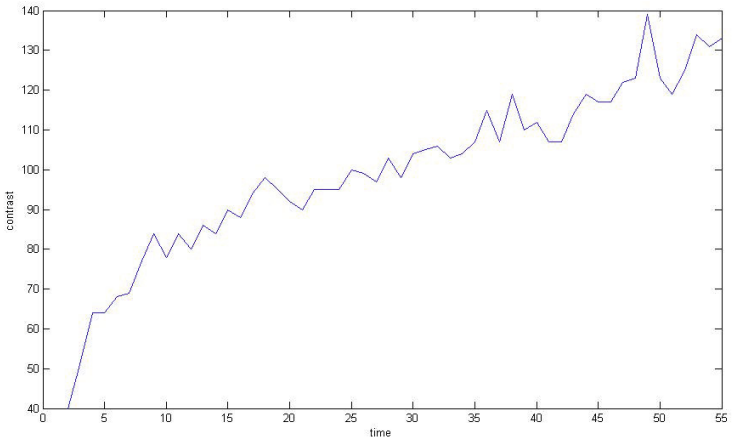
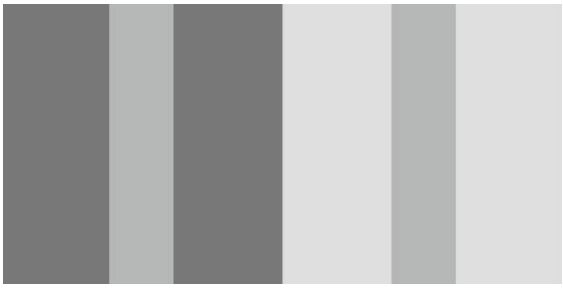
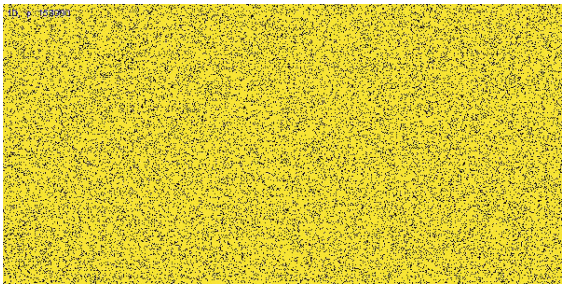


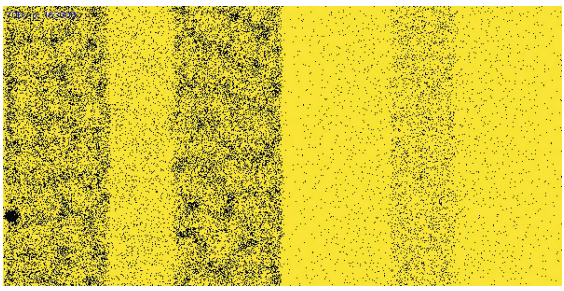
Fig. 12 Plot of the emergence of global contrast by particle flux in response to the Chevreul staircase stimulus (difference in population density range, from an initially random density distribution, to halting at 550 scheduler steps)



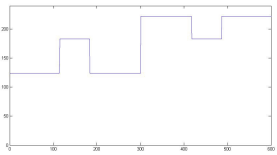
(a) SBC image



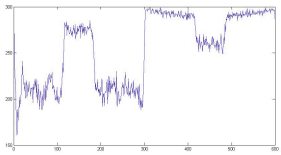
(b) t=10



(c) t=700



(d) original section



(e) population density t=700

Fig. 13 Collective representation of the Simultaneous Brightness Contrast (SBC) illusion: (a) Original greyscale Image of two large squares, each with a central band of identically light grey, (b) initial uniform population distribution of particle population, (c) population distribution after 700 scheduler steps showing increased population density at lighter (greater attractant concentration) regions and greater density in the left central grey strip, (d) cross-section profile of original image stimulus, (e) cross-section of population density at t=700 showing illusory percept of the left central grey strip as brighter than the right central grey strip

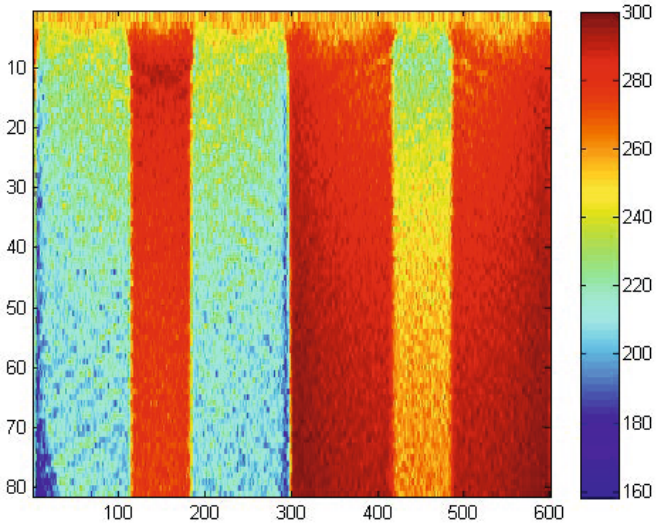


Fig. 14 Space-time plot of population density flux under Simultaneous Brightness Contrast stimulus, time proceeds downwards

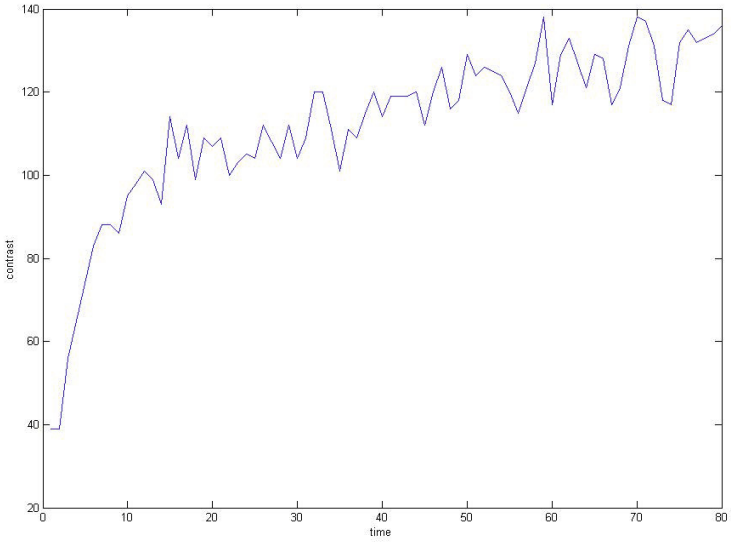


Fig. 15 Plot of the emergence of global contrast over time by particle flux in response to the SBC stimulus (difference in population density range, from an initially random density distribution, to halting at 800 scheduler steps)

by the attractant stimuli and particles migrated to regions of higher concentration. The flux of particles changed the population density, with higher occupancy emerging in regions which corresponded to lighter areas of the original image (Fig. 10c). Note that within each bar the density is greater towards the left side of the bar. This is caused by influx from the darker bar to the left and efflux towards lighter bars to the right. A plot of the population density at 400 scheduler steps demonstrates both the scalloped response to each bar junction and the global increase in density towards the lighter bars (corresponding to the perceived lightness of the global image, Fig. 10e). A space-time plot of the evolution of population density indicates that the scalloped effect occurs immediately after presentation with the stimuli (Fig. 11). The differences in density between the bars also starts immediately after presentation but the full increase in 'brightness' between each bar occurs much later, due to the time taken for particles to migrate towards higher concentration areas (Fig. 10e, note the gradual increase in density in each bar over time). The increase in contrast of the global (entire arena) population density over time can be seen in Fig. 12. This contrast is caused by the relatively slow flux of particles across the entire arena and corresponds to a coarse representation of the image brightness.

The Chevreul staircase consists of contiguous regions of gradually increasing lightness. Another illusory perception of lightness occurs in regions with non-contiguous increases in lightness. This is known as the Simultaneous Brightness Contrast (SBC) effect.[1]

A simple example of SBC is shown in Fig. 13a. The image consists of two large squares, the left-most square in dark grey and the right-most square in a lighter grey. Overlaying the centre of each square is a vertical band of grey (intermediate in lightness between the left and right squares). Although the vertical grey bands are of equal lightness (see cross-section in 13d), their lightness is perceived differently: the band on the left is typically perceived as lighter than the band on the right. As with the Chevreul staircase the explanatory mechanisms for this illusion have been suggested as LI at the retinal and cortical level. How does the virtual plasmodium respond when presented with this stimuli?

We initialised the model on a 600×300 lattice with 153,000 particles (approximately the same density as used in the Chevreul staircase experiments) with uniform initial distribution (13b) and periodic boundary conditions. The lightness differences of the SBC image areas were represented as differing attractant concentration profiles. The population response is to migrate from areas of low concentration (brightness) towards areas of high concentration. This flux is mobilised at the junctions between light and dark regions and a global difference in population density is produced (13c). Because the vertical grey band in the darker square is surrounded by a darker region, it receives influx from the darker square. The vertical band within the lighter square is surrounded by a lighter region and there is an efflux of particles to the lighter square. Simultaneously there is transport of particles at the centre of the image (the border between the two large squares) towards the lighter side, and also

[1] Note that the terms *lightness* (SLC) and *brightness* (SBC) are both used in the literature to describe the same perceptual effect, though their actual definitions are somewhat different, depending on the particular presentation of the stimuli.

at the left and right edges of the image (due to periodic boundary conditions). The resulting population density plot (13e) shows that the vertical band in the left square is perceived as 'brighter' than the band in the right square (although the stimuli concentration presented was identical). This illusory percept matches that of human subjects when presented with the same stimulus type [6].

A space-time plot of the evolution of population density demonstrates the differences in population density emerging as the experiment progresses. As with the Chevreul staircase, the local response to adjacent stimuli in the SBC image is instantaneous, whereas the global response to overall brightness changes in the image takes significantly longer, (also seen in the contrast plot in Fig. 15), again caused by the relatively slow movement of particles across the image areas.

7 Conclusions

We have demonstrated the results of modelling experiments into the generation of spatial contrast enhancement analogous to Lateral Inhibition in unorganised non-neural systems using a multi-agent model of slime mould *Physarum polycephalum*. The results show the classic LI contrast enhancement in response to attractant stimuli and its opposite counterpart behaviour (Lateral Activation) in response to adverse stimuli (simulated light irradiation). These effects does not require pre-existing inhibitory connectivity and are generated by bulk transport of the particles comprising the virtual material, initiated at the borders of stimuli projection. Restoration of uniform baseline activity (population density distribution) is established when stimuli are removed. In addition to the local edge contrast enhancement we observed long-term changes in population density distribution when the population was presented with a more complex attractant stimulus pattern. This was caused by the flux of particles towards brighter regions of the image and corresponds to a (rather crude) collective response to the global brightness of the original image stimuli, including the scalloped intensity profile of the Chevreul staircase and the perceived difference of two identically bright patches in the Simultaneous Brightness Contrast (SBC) effect. Interestingly, this simple mechanism reproduces the illusory human percepts in both the Chevreul staircase and SBC figures.

How may these simple mechanisms correspond to real-life perception in neural systems? It is possible to relate the response of the model to attractant and repellent stimuli to the response of different populations (ON or OFF type respectively) of bipolar cells in the retina of the human visual system. In the model, however, we only have a single 'type' of particle and the opposing response to attractant and repellent stimuli loosely correspond to the response to light and dark stimuli in the retina. The receptive field around retinal stimuli is approximated by the region of influx of efflux of particles from the local stimuli and the representation of the *eigengrau* level (the baseline activity in the absence of any stimulus) may be approximated by the population density of the particles when no attractant stimuli or repellent stimuli are present.

The most notable feature of this approach is that LI phenomena can be approximated without explicit fixed inhibitory connections. This suggests possible mechanisms by which simple organisms without neural tissue may achieve sensory contrast enhancement. In organisms such as slime mould the internal protoplasmic transport of cellular material could be harnessed to generate the LI mechanism. The sensory contrast enhancement afforded by LI could allow for the enhancement of weak spatial stimuli (such as nutrient location). Conversely, the LA mechanism would amplify weak hazardous stimuli, providing alternate migration paths away from hazardous regions. The relatively slow restoration of baseline activity also allows a temporary memory effect denoting the approximate location of attractant and hazard stimuli. Because the LI and LA phenomena in this model do not rely on fixed inhibitory connectivity it is particularly suited to systems and organisms which have adaptive architectures and body plans respectively.

In the context of adaptive materials and robotics applications the mechanism illustrates how complex sensory behaviour can be distributed within an unorganised material (or robotic collective) itself. This allows greater freedom from having to pre-specify connectivity to implement sensory contrast enhancement and allows redundancy for individual faulty components. We hope that ongoing research may lead to other unorganised material approximations of complex neural functions seen in brightness perception (including illusory phenomena such as neon colour spreading, illusory contours and brightness assimilation effects), and implementation of other spatial feature detectors (including orientation detection, edge completion, gestalt phenomena, optic flow), and direction discrimination.

Acknowledgements. This work was supported by the EU research project "Physarum Chip: Growing Computers from Slime Mould" (FP7 ICT Ref 316366).

References

1. Adamatzky, A.: Physarum Machine. Implementation of a Kolmogorov-Uspensky machine on a biological substrate. Parallel Processing Letters 17(4), 455–467 (2007)
2. Adamatzky, A.: Developing proximity graphs by Physarum polycephalum: does the plasmodium follow the toussaint hierarchy. Parallel Processing Letters 19, 105–127 (2008)
3. Adamatzky, A., Jones, J.: Towards Physarum robots: computing and manipulating on water surface. Journal of Bionic Engineering 5(4), 348–357 (2008)
4. Adamatzky, A., Jones, J.: Programmable reconfiguration of. Physarum machines. Natural Computing 9(1), 219–237 (2010)
5. Aono, M., Hara, M.: Amoeba-based nonequilibrium neurocomputer utilizing fluctuations and instability. In: Akl, S.G., Calude, C.S., Dinneen, M.J., Rozenberg, G., Wareham, H.T. (eds.) UC 2007. LNCS, vol. 4618, pp. 41–54. Springer, Heidelberg (2007)
6. Blakeslee, B., McCourt, M.E.: A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. Vision Research 44(21), 2483–2503 (2004)
7. Gale, E., Adamatzky, A., de Lacy Costello, B.: Slime mould memristors. BioNanoScience, pp. 1–8 (2013)
8. Hartline, H.K., Ratliff, F.: Inhibitory interaction in the retina of limulus. In: Physiology of Photoreceptor Organs, pp. 381–447. Springer (1972)
9. Houtgast, T.: Psychophysical evidence for lateral inhibition in hearing. The Journal of the Acoustical Society of America 51(6B), 1885–1894 (1972)

10. Ishiguro, A., Shimizu, M., Kawakatsu, T.: A modular robot that exhibits amoebic locomotion. Robotics and Autonomous Systems 54(8), 641–650 (2006)
11. Jones, J.: Characteristics of pattern formation and evolution in approximations of Physarum transport networks. Artificial Life 16(2), 127–153 (2010)
12. Jones, J.: The emergence and dynamical evolution of complex transport networks from simple low-level behaviours. International Journal of Unconventional Computing 6(2), 125–144 (2010)
13. Jones, J., Adamatzky, A.: Emergence of self-organized amoeboid movement in a multi-agent approximation of. Bioinspiration and Biomimetics 7(1), 016,009 (2012), http://stacks.iop.org/1748-3190/7/i=1/a=016009
14. Kandel, E.R., Schwartz, J.H., Jessell, T.M.: Principles of neural science, vol, vol. 4. McGraw-Hill, New York (2000)
15. Macknik, S.L., Martinez-Conde, S.: The spatial and temporal effects of lateral inhibitory networks and their relevance to the visibility of spatiotemporal edges. Neurocomputing 58, 775–782 (2004)
16. Nakagaki, T., Kobayashi, R., Nishiura, Y., Ueda, T.: Obtaining multiple separate food sources: behavioural intelligence in the *Physarum* plasmodium. R. Soc. Proc. Biol. Sci. 271(1554), 2305–2310 (2004)
17. Nakagaki, T., Yamada, H., Toth, A.: Intelligence: Maze-solving by an amoeboid organism. Nature 407, 470 (2000)
18. Peromaa, T.L., Laurinen, P.I.: Separation of edge detection and brightness perception. Vision Research 44(16), 1919–1925 (2004)
19. Pershin, Y., La Fontaine, S., Di Ventra, M.: Memristive model of amoeba learning. Physical Review E 80(2), 021,926 (2009)
20. Pessoa, L., Mingolla, E., Neumann, H.: A contrast-and luminance-driven multiscale network model of brightness perception. Vision Research 35(15), 2201–2223 (1995)
21. Sakiyama, T., Gunji, Y.P.: The müller-lyer illusion in ant foraging. PloS One 8(12), e81,714 (2013)
22. Serino, A., Haggard, P.: Touch and the body. Neuroscience & Biobehavioral Reviews 34(2), 224–236 (2010)
23. Shirakawa, T., Adamatzky, A., Gunji, Y.P., Miyake, Y.: On simultaneous construction of voronoi diagram and delaunay triangulation by Physarum polycephalum. International Journal of Bifurcation and Chaos 19(9), 3109–3117 (2009)
24. Takagi, S., Ueda, T.: Emergence and transitions of dynamic patterns of thickness oscillation of the plasmodium of the true slime mold Physarum polycephalum. Physica D 237, 420–427 (2008)
25. Takamatsu, A.: Spontaneous switching among multiple spatio-temporal patterns in three-oscillator systems constructed with oscillatory cells of true slime mold. Physica D: Nonlinear Phenomena 223(2), 180–188 (2006)
26. Tani, I., Yamachiyo, M., Shirakawa, T., Gunji, Y.: Kanizsa illusory contours appearing in the plasmodium pattern of physarum polycephalum. Frontiers in Cellular and Infection Microbiology 4 (2014)
27. Tsuda, S., Aono, M., Gunji, Y.P.: Robust and emergent *Physarum* logical-computing. BioSystems 73, 45–55 (2004)
28. Tsuda, S., Zauner, K.P., Gunji, Y.P.: Robot control with biological cells. BioSystems 87, 215–223 (2007)
29. Urban, N.N.: Lateral inhibition in the olfactory bulb and in olfaction. Physiology & Behavior 77(4), 607–612 (2002)

Biomimicry of Crowd Evacuation with a Slime Mould Cellular Automaton Model

Vicky S. Kalogeiton, Dim P. Papadopoulos, Ioannis P. Georgilas, Georgios Ch. Sirakoulis, and Andrew I. Adamatzky

Abstract. Evacuation is an imminent movement of people away from sources of danger. Evacuation in highly structured environments, e.g. building, requires advance planning and large-scale control. Finding a shortest path towards exit is a key for the prompt successful evacuation. Slime mould *Physarum polycephalum* is proven to be an efficient path solver: the living slime mould calculates optimal paths towards sources of attractants yet maximizes distances from repellents. The search strategy implemented by the slime mould is straightforward yet efficient. The slime mould develops may active traveling zones, or pseudopodia, which propagates along different, alternative, routes the pseudopodia close to the target loci became dominating and the pseudopodia propagating along less optimal routes decease. We adopt the slime mould's strategy in a Cellular-Automaton (CA) model of a crowd evacuation. CA are massive-parallel computation tool capable for mimicking the Physarum's behaviour. The model accounts for Physarum foraging process, the food diffusion, the organism's growth, the creation of tubes for each organism, the selection of optimum path for each human and imitation movement of all humans at each time step towards near exit. To test the efficiency and robustness of the proposed CA model, several simulation scenarios were proposed proving that the model succeeds to reproduce sufficiently the Physarum's inspiring behaviour.

1 Introduction

The need for realistic and efficient in case of emergency crowd dynamics modeling approaches has been considered a scientific topic of great importance. Having in

Vicky S. Kalogeiton · Dim P. Papadopoulos · Georgios Ch. Sirakoulis
Department of Electrical and Computer Engineering,
Democritus University of Thrace, Greece
e-mail: gsirak@ee.duth.gr

Ioannis P. Georgilas · Andrew I. Adamatzky
Centre for Unconventional Computing, University of the West of England, Bristol, UK

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*,
Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_7

mind that our safety and comfort depend crucially on our fellow crowd members and on the design and operation of the facility we are in, the quest of proper computational tools for modeling crowd behaviour is continuous.In the past, the main constraint in deep studying of crowd behaviour was the large number of required computations and the respective limitation of computational resources. Crowd was modeled as homogeneous mass characterised by properties of a moving fluid. In specific, pedestrian movement simulation models can mainly be categorized into macroscopic and microscopic ones. Macroscopic models focus on the total number of the members of the crowd. In other words, they study the major characteristics of the flow of the individuals rather than the features of the individuals themselves. Microscopic models study the attitude of each of the individuals and their interaction with the other members of the crowd. Furthermore, microscopic models describe the spatial and temporal behavior of each individual. The response of such models is very interesting, since they manage to represent virtual crowds to realistic autonomous behaviors. Cellular Automata (CA) [62] models, agent–based model and social–force models belong to this category [8, 11, 40, 18, 67, 49, 61].

While particle and CA based approaches mostly aim at generating quantitative results about pedestrian and crowd movement, agent based models sometimes aim at the generation of effective visualizations of believable pedestrian dynamics, and therefore the above approaches do not necessarily share the same notion of realism and validation [61]. Works like [8, 22] separate the pedestrians from the environment and grant them with a complex behavioral specification. Other approaches ([36, 45]) aim at generating visually effective and believable pedestrians and crowds in virtual worlds [45]. Finally, works like ([41]), employ cognitive agent models for different goals, but they are not generally focused on making predictions about pedestrian movement for sake of decision support. Social–force models describe the behavior of an individual through social sectors (virtual physical forces) that arise from the social attitude of the characters. Taking into consideration virtual social forces equivalent to real ones, such as repulsive interaction, friction forces and some fluctuations, they enable the resolution of Newton's motion equations of individuals. Compared to other models, social–force ones describe the attitude of an individual more realistically. Each motion parameter is assigned to physical meaning, it is unique for each of the members and its selection is often based on some empirical data. Social forces form the behavior of the members of the crowd through a combination of social–psychological and natural factors [20].

More specifically, Helbing *et al.* used repulsive and attractive forces to simulate the interaction among individuals and obstacles, offering a realistic behavior of individuals' prompting and variable percentages of flow as well. Furthermore, literature includes several studies and applications that use simulation approaches of particles for low–dense crowd formations. Specifically, Brogan and Hodgins [10] employed systems of particles and potentials in order to depict the motion of special–structured groups of people. Musse extended the model of social–forces, introducing the notion of individualism [9].

In CA based approach [62], the space under study is presented as a unified grid of cells with local attributes, which are generated by a set of rules that describe

the behavior of the individuals [12]. The state of each cell is defined according to the rules, the state of the cell at the previous time step and the current state of neighboring cells. Recent approaches suggest that pedestrian crowd consists of discrete individuals, able to react with their surroundings. This requires a large number of computations. A possible solution could be modern computer power combined with the use of a CA computational model that is featured by massive parallelism. CA are very effective in simulating physical systems and solving scientific problems [49]. They can capture the essential features of systems where global behavior arises from the collective effect of simple components, which interact locally [15]. Furthermore, evacuation process is inherently complex, i.e. a system is multi–parameterized and its response cannot be easily estimated. There are interactions among occupants, places and environment and there are also socio–psychological parameters that should be taken under consideration [32, 66]. Evacuation could be defined as a non linear problem with many factors affecting it. Literature records many CA–based models investigating crowd behavior under various circumstances. The impact of environmental conditions [14, 42, 59, 67] and bi–directional pedestrian behavior [63, 25] has been examined. Interactions among pedestrians, friction effects [31, 43]. and herding behavior [40, 18] have also been considered. Furthermore, CA originated models are reported that focus on human behaviors, such as inertial effects, unadventurous effect and group effect [68]. Some models treat pedestrians as particles subject to long–range forces [7] and others use walkers leaving a trace by modifying the underground on their paths [11].

Macroscopic models focus on specific attitudinal attributes of separate members as well as on a generic handling of the pedestrians as a flow of individuals. Such models are very important in evaluating and predicting the needs of vehicle traffic or studying the dynamics of large scale constructions. In this category belong the models of regression that make use of statistically established relations on flow variables, in order to predict flow procedures of individuals under certain circumstances [35]. The models of route choice describe the path searching based on the concept of practicality. The members choose their destination aiming at the maximizing of the benefits of their route (comfort, time of the trip, etc.) [23]. Anticipation models use Markov chain models to describe the way that crowd move from one node of the net to the other. The gas motion models use a proportion based on the fluid or gas dynamics as well as on partial differential equations, in order to describe how the density and the velocity of the crowd change with time [21].

Instead of using a single yet efficient model for simulation of the evacuation process, here the idea of mimicking the physarum behavior during its foraging process in order to get closer to food sources is analysed. In more details, the touchstone of our inspiration arrives from the fact that physarum as a real biological organism enables itself to find one of the best/shortest possible ways towards food in an almost complex yet well described process [2].

Thus, in [3] we shown in laboratory experiments and numerical simulation that if plasmodium (vegetative form) of Physarum is inoculated in a maze's peripheral channel and an oat flake (source of attractants) is placed in the maze's central chamber then the plasmodium grows toward the target oat flake and connects the flake

with the site of original inoculation with a pronounced protoplasmic tube. The protoplasmic tube represents a path in the maze. The plasmodium solves maze in one pass because it is assisted by a gradient of chemo–attractants propagating from the target oat flake.

The driving force to do so arrives mainly from the need of self–preservation. Taking into account that in some cases and beyond quantum processes, the same need, i.e. self–preservation can be found typically in a macroscopic level to the humans similar to the physarum behaviour we are triggered to explore possible applications of the aforementioned phenomenological behaviour to advance our crowd evacuation models. Although such a so–called large scale phenomenon, i.e. human crowd evacuation is supposed to be hardly predictable, it tends to present some interesting features that worth further researching. Having in mind that the Physarum, beyond the simplicity of its plasmodium's body and the fact that it is extremely easy to cultivate and handle, has demonstrated complicated and robust computing capacity when it confronted solving mazes problems [37]. Due to this capacity it can be considered as a fine candidate for modelling in an unconventional way the crowd evacuation process. Nevertheless, real time experiments with Physarum are time consuming; thus, is apparent the necessity of modelling its behaviour and the emergent computation abilities. Towards this direction, some slime mould models have been already proposed granting optimistic results. In specific, Tero and Nakagaki [54] introduced a mathematical model for mimicking of the true slime mold adaptive dynamics while constructing tubular networks. A multi–agent model was proposed by Jones *et al.* [26], while some Cellular Automata (CA) based models were proposed by Gunji *et al.*, who showed that their latticeagents–based model, namely CELL, is moving like an amoeba [19], as well as by Liu *et al.* who introduced an enhanced version of the agents–based CELL [33], by Tsompanas and Sirakoulis, who presented a CA model to simulate the plasmodiums foraging strategy and formation of a tubular network [57] and more recently by Kalogeiton *et al.* proposing a CA model for slime mould sufficient enough to result in successful Simultaneous Localisation and Mapping (SLAM) for robotic applications [27].

In this chapter by bio–mimicking Physarum's behaviour, a CA based crowd evacuation model is described in detail. In particular, the proposed CA model takes into account while mimicking the Physarum foraging process, the food diffusion, the organism's diffusion, the creation of tubes for each organism, the selection of optimum tube for each human in correspondence to the under study crowd evacuation and finally the movement of all humans at each time step towards near exit. All these procedures are repeated during the evolution of the CA model, as many times needed until all humans exit the under study indoor environment. It should be mentioned that in the proposed model all the characteristics of the space as well as of the humans' movement are considered based on the related pedestrian dynamics literature. To test the efficiency and the robustness of the presented CA model several simulation scenarios were proposed. More specifically, five different experimental results that indicate and prove the integrity of the algorithm were introduced. In all the examined cases the efficiency of the proposed evacuation model was sufficiently indicated in qualitative manner by the existence of basic crowd characteristics, such

as clustering effects, the buildup of pressure, congestion at the exits, clogging effects at bottlenecks, jamming at widening and arching effects, which became apparent in most of the under study environments. All these results confirm the effectiveness of the proposed model for human evacuation scenarios.

The paper is structured as follows. In Section 2 some preliminaries on the *P. polycephalum* organism and CA basic theory are presented. Moreover, the preliminaries and the basic assumptions of the proposed here CA crowd evacuation model are given in Section 3, while the same Section 3 outlines the proposed here model in detail enlighten all the metaphors and simplifications needed to transform the Physarum foraging process to a efficient CA based crowd evacuation model. The results of the aforementioned model applied in fivedifferent simulation environments are depicted in Section 4. Finally, conclusions are drawn in Section 5.

2 Preliminaries: Physarum and CA Theory

2.1 Physarum

Plasmodium is a vegetative stage of the acellular slime mould *P. polycephalum*, a syncytium, that is, a single cell with many nuclei, which feeds on microscopic particles [53]. When foraging for its food the plasmodium propagates towards sources of food particles, surrounds them, secretes enzymes and digests the food. Typically, the plasmodium forms a congregation of protoplasm covering the food source. When several sources of nutrients are scattered in the plasmodium's range, the plasmodium forms a network of protoplasmic tubes connecting the masses of protoplasm at the food sources.

The plasmodium is a unique user–friendly biological substrate from which experimental prototypes of massive–parallel amorphous biological computers are designed [2]. During its foraging behaviour the plasmodium spans scattered sources of nutrients with a network of protoplasmic tubes. The protoplasmic network is optimised to cover all sources of food and to provide a robust and speedy transportation of nutrients and metabolites in the plasmodium body. The plasmodium's foraging behaviour can be interpreted as computation. Data are represented by spatial configurations of attractants and repellents, and results of computation by structures of a protoplasmic network formed by the plasmodium on the data sets [39, 38, 2]. The problems solved by plasmodium of *P. polycephalum* include shortest path [39, 38], implementation of storage modification machines [1], Voronoi diagram [46], Delaunay triangulation [2], logical computing [58], and process algebra [44]; see overview in [2].

In [4] we conducted pioneer laboratory experiments with Nylon terrains of USA and Germany. We used the Physarum to approximate route 20, the longest road in USA, and autobahn 7, the longest national motorway in Europe. We found that Physarum builds longer transport routes on 3D terrains, comparing to flat substrates

yet sufficiently approximates man–made transport routes studied. We demonstrate that nutrients placed in destination sites affect performance of Physarum, show how the Physarum navigates around elevations. In cellular automaton models of the Physarum we shown variability of the protoplasmic routes might depends on physiological states of the slime mould. Results presented will contribute towards development of novel algorithms for sensorial fusion, information processing, and decision making, and will provide inspirations in design of bio–inspired amorphous robotic devices.

Our results [4] demonstrated that Physarum is capable of solving practical tasks of path–finding even when confined to a non–friendly substrate of bare plastic. Thus we decided to test the Physarum's abilities on imitating tasks of evacuation in a physical model of a building.

2.2 CA Theory

Cellular automata (CA) are models of physical systems, where space and time are discrete and interactions are local [13]. In brief, CA were originally proposed by John von Neumann (1966), as formal models of self–reproducing organisms, who was thinking of imitating the behaviour of a human brain in order to build a machine able to solve very complex problems [62]. His ambitious project was to show that complex phenomena can, in principle, be reduced to the dynamics of many identical, very simple primitives, capable of interacting and maintaining their identity [13]. Following a suggestion by Ulam [1952], von Neumann adopted a fully discrete approach, in which space, time and even the dynamical variables were defined to be discrete. Consequently, CA are very effective in simulating physical systems and solving scientific problems, because they can capture the essential features of systems where global behaviour arises from the collective effect of simple components which interact locally [65].

A CA consists of a regular uniform n-dimensional lattice (or array), usually of infinite extent. At each site of the lattice (cell), a physical quantity takes on values. The value of this physical quantity over all the cells is the global state of the CA, whereas the value of this quantity at each site is its local state. Each cell is restricted to local neighborhood interaction only and, as a result, it is incapable of immediate global communication [62].

A CA is characterized by five properties:

1. The number of spatial dimensions (n).
2. The width of each side of the array (w). w_j is the width of the j^{th} side of the array, where $j = 1,2,3,...,n$.
3. The width of the neighborhood of the cell (r).
4. The states of the CA cells.
5. The CA rule, which is an arbitrary function F.

The state of a cell, at time step $(t+1)$, is computed according to F, a function of the state of this cell at time step (t) and the states of the cells in its neighborhood at time step (t). The neighborhood of each cell is defined by variable N. For a 2-d CA, two neighborhoods are often considered: Von Neumann and Moore neighborhoor, respectively. Equation 1 defines the Von Neumann neighborhood of range r.

$$N^N_{(x_0,y_0)} = \{(x,y) : |x-x_0| + |y-y_0| \leq (r)\}. \quad (1)$$

For a given cell (x_0, y_0) and range r, Moore neighborhood can be defined by the following formula:

$$N^M_{(x_0,y_0)} = \{(x,y) : |x-x_0| \leq (r), |y-y_0| \leq (r)\}. \quad (2)$$

In most practical applications, when simulating a CA rule, it is impossible to deal with an infinite lattice. The system must be finite and have boundaries. Clearly, a site belonging to the lattice boundary does not have the same neighborhood as other internal sites. In order to define the behavior of these sites, neighborhood is extending for the sites at the boundary, thus leading to various types of boundary conditions such as periodic (or cyclic), fixed, adiabatic or reflection.

CA have sufficient expressive dynamics to represent phenomena of arbitrary complexity [65, 16, 29, 50] and at the same time can be simulated exactly by digital computers, because of their intrinsic discreteness, i.e. the topology of the simulated object is reproduced in the simulating device [60]. The CA approach is consistent with the modern notion of unified space-time. In computer science, space corresponds to memory and time to processing unit. In CA, memory (CA cell state) and processing unit (CA local rule) are inseparably related to a CA cell [51, 48]. Furthermore, they can easily handle complicated boundary and initial conditions, inhomogeneities and anisotropies [51]. In addition, algorithms based on CA run quickly on digital computers [56]. Models based on CA lead to algorithms, which are fast when implemented on serial computers, because they exploit the inherent parallelism of the CA structure. These algorithms are also appropriate for implementation on massively parallel computers [52], such as the Cellular Automaton Machine (CAM) [64] or Field Programmable Gate Arrays (FPGAs) [17, 34, 24].

3 Proposed Bio–inspired Evacuation Method

As mentioned before, the method that is proposed in this paper refers to the design and implementation of a novel bio–inspired/bio–mimicking model for human evacuation scenarios. This method is based on the plasmodium of *P. polycephalum* coupled with the CA parallel computational tool.

In particular, it is attempted to simulate the movement of a crowd, when the crowd is located in an indoor environment and is trying to evacuate it. Each human located in the environment is represented by plasmodium of Physarum, acquiring the

functionalities the organism has under specific circumstances. This particular microorganism is deployed into a nutrient–poor substrate, which represents the indoor to–be–evacuated environment. The position of each organism is the corresponding position of each individual human, while at the same time all exits of the environment are considered as food sources for the Physarum. In that way, all plasmodiums are spreading towards food source and eventually construct a tubular network, the cytoplasm, which connects optimally each plasmodium with the food sources, i.e. each human with the exits of the space. Therefore, for each human the optimum and shortest path/ tube is selected and is the one that the human follows in order to exit the environment. In case the path, which a human follows, is interrupted, meaning that the human is not able to move since the target position is occupied by another human, the plasmodium is required to re–diffuse and hence, to re–create a network based on the new conditions. This procedure results in the successful evacuation of indoor environments.

3.1 Space Delineation

All evacuation scenarios examined in this paper refer to indoor environments. Thus, it is important to determine the space in which the evacuation takes place.

According to CA the space examined in each case is divided to cells, thus resulting in a CA grid. Each cell is individual and independent not to mention the fact that it is considered as possible position (for humans, obstacles, exits etc.). The dimension of each cell is crucial, since it should be large enough to represent a human and small enough compared to the total space. According to recent studies demonstrated in [11], each human can cover a minimum area of: 40×40 cm^2. This dimension has been chosen, since according to the corresponding studies the proposed area represents the region a human needs when standing in order to feel comfortable with the environment and at the same time in order not to be possible for another human to occupy the same space. The basic definitions that refer to the under study space are summed up as follows:

- Assume that the space Ω is rectangular, without loss of generality, with dimensions: $L_1 \times L_2$ m^2.
- The space Ω is partitioned into computational cells $c_{i,j}$, where $c_{i,j} = 1,\ldots,K$, such that $c_{i,j} \cap c_{m,n} = \oslash, (i,j) \neq (m,n)$ and $\bigcup_{i,j=1}^{K} c_{i,j} = \Omega$.
- The number of cells $c_{i,j}$ of each side of the under study space is $N_{1c_{i,j}} = \frac{L_1}{0.4}$ and $N_{2c_{i,j}} = \frac{L_2}{0.4}$, respectively.
- Thus, it holds that: $i = 1,\ldots,N_{1c_{i,j}}$ and $j = 1,\ldots,N_{2c_{i,j}}$.
- Define the total number of cells as: $N_{c_{i,j}} = N_{1c_{i,j}} \times N_{2c_{i,j}}$.

3.2 Initial Declarations

As mentioned before, the main goal of this method is the simulation of the human evacuation process. To begin with, it is important to define the matching rules between the real environment and its corresponding biological one. The elements needed in order to design an evacuation model could be summed up as the position of each human inside the space, the position of the exits and of the obstacles into the environment and of course, the remaining ”free space”, meaning the space that is not occupied neither by humans nor by obstacles/exits.

Thus, the matching of each of these elements with their corresponding biological ones derives from the human ”representation”. Each human is represented by the Physarum, and it is assumed that an initial quantity of the organism in plasmodium state is placed - and hence accumulated - in the particular cell that is occupied by the corresponding human. As a result, each human is represented by a plasmodium while in the rest of the space (including both the free space and the obstacles) no organism exists. As far as the exits of the environments is concerned, these are represented by the food sources, meaning that for each cell an exit exists, a food source is placed.

At this point, it should be mentioned the environment circumstances under which the Physarum will function. According to [47] in case the plasmodia of Physarum develop in a nutrient–rich substrate (agar gel containing crushed oat flakes) the colliding protoplasmic fronts merge. On the other hand, when the plasmodia develop on a nutrient-free substrate, protoplasmic tubes are formed between the loci of the concentrated cytoplasm and the food sources. Figure 1, which follows, illustrates the development of plasmodia on a plain substrate during time, proving the formation of the tubular network between the food sources. This image is based on real biological experiments [47].

The formation of the tubular network constitutes the main reason for having chosen this microorganism to represent the human. In an evacuation scenario each human follows a path to reach a building exit. The ability of the Physarum to find a shortest path creating the tubes is the most profound feature exploited by this method. That means that the organism is able to find the shortest path between its position and the exit carving in that way the route that the human will follow. It should be noted that, the environment that each cytoplasm is developed, is considered to be nutrient–free.

In this nutrient–free substrate the microorganism can be developed only under appropriate circumstances. Hence, the coordinates of the inaccessible by humans regions of the real environment (e.g. the obstacles, the walls etc.) indicate to the corresponding coordinates of the assumed substrate, in which physical barriers meet. The term physical barrier refers mainly to constrained circumstances under which the plasmodium is not able to develop, such as under light regime [55] or next to grains of salt [5].Thus, the representation of obstacles and walls could be done by using light. Figure 2 demonstrates an example of how Physarum’s plasmodium responds to light.

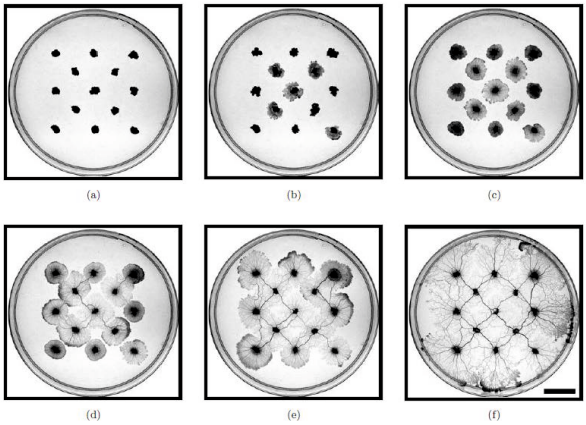


Fig. 1 Development of plasmodia on a nutrient-free substrate. The Physarum is concentrated in different locations in the substrate - shown by black color in (a) - and then the network is formulated. Each figure indicates the state of the plasmodia (a) 12 h, (b) 23 h, (c) 34 h, (d) 47 h, (e) 69 h, (f) 80 h after the inoculation on the substrate. (d) and (e) When the plasmodia come to contact, they show avoidance of the other plasmodia, though they fuse at some points, in a very limited manner. (f) Tube formation proceeded with time, and the enhanced tubular network is constructed. (Scale Bar 2 cm) [47]

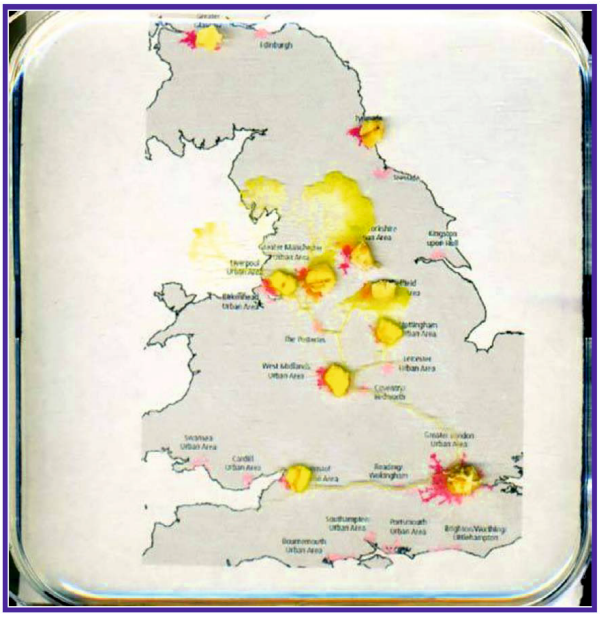


Fig. 2 The developing Physarum network adopted by [5]. Typical experimental setup of UK transportation network where ten most populous urban areas are represented by oat flakes, plasmodium is inoculated in London, the plasmodium spans oat flakes by protoplasmic transport network.

3.3 Method Overview

After the determination of the basic contents of the method, a detailed description of the proposed bio–inspired evacuation algorithm follows. The initial steps are the food diffusion and the organism growth. The diffusion of the food as well of the microorganism takes place based on the proposed here CA model.

Chemo–attractant Diffusion from Sources of Nutrients. After defining earlier the space Ω and the number $N_{c_{i,j}}$ of cells $c_{i,j}$, assume that:

- In the $N_{c_{i,j}}$ cells there are K active cells, such that $K \subseteq N_{c_{i,j}}$.
- The set of active cells is defined as $A = \left\{c_{i,j}^{A_1}, c_{i,j}^{A_2}, c_{i,j}^{A_3}, \ldots, c_{i,j}^{A_K}\right\}$.
- The state of these K active cells is denoted by $p_{i,j}$.

Active cells are the cells that the plasmodium is able to diffuse and therefore to grow cytoplasm. These cells correspond to the so called free space of the environment.

- The set of repellents $R = \{R_1, R_2, R_3, \ldots\}$ is defined as the total of cells that the plasmodium cannot reach.
- Let r be the number of repellents R, so that: $R = \{R_1, R_2, R_3, \ldots, R_r\}$.
- Define $\left\{c_{i,j}^{R_1}, c_{i,j}^{R_2}, c_{i,j}^{R_3}, \ldots, c_{i,j}^{R_r}\right\}$ the corresponding cells of the set of repellents $R = \{R_1, R_2, R_3, \ldots, R_r\}$.

Repellents are are physical stimuli that cause an organism to migrate away or react adversely. It is known that the plasmodium of Physarum avoids light and salt. Thus, domains of high illumination together with salt concentrations are repellents such that each repellent R is characterised by its position and intensity of illumination or force of repelling.

- The set of attractants $F = \{F_1, F_2, F_3, \ldots\}$ are sources of nutrients on which the plasmodium feeds.
- Let s be the number of food sources F, so $F = \{F_1, F_2, F_3, \ldots, F_s\}$.
- Let $\left\{c_{i,j}^{F_1}, c_{i,j}^{F_2}, c_{i,j}^{F_3}, \ldots, c_{i,j}^{F_r}\right\}$ be the corresponding cells of the set of food sources $F = \{F_1, F_2, F_3, \ldots, F_s\}$.

It is still the subject of discussion on how exactly plasmodium feels the presence of attracts. Indeed diffusion of some kind is involved [6]. Thus, the representation of the diffusion process both of the food and the plasmodium takes place based on the proposed here CA model. The neighbourhood chosen for both cases is the Moore neighbourhood, since according to [28, 30] is appropriate and capable of simulating the diffusion of a liquid and of a smell.

The food diffusion takes place only to the K active cells, given that the smell can pass through a human, which occupies a cell, but not through a wall. The smell, $Smell_{i,j}$, starts from the cells $\{c_{i,j}^{F_1}, c_{i,j}^{F_2}, c_{i,j}^{F_3}, \ldots\}$, that represent the exits, namely the food sources, and spreads at each time step t to their neighbour cells according to Equation 3:

$$Smell^{t+1}_{(i,j)} = Smell^{t}_{(i,j)} + s_1 \times (SVN) + s_2 \times (RMN), \tag{3}$$

where *SVN* expresses the *von Neumann* neighbourhood and it is defined as follows:

$$\begin{aligned} SVN = Smell^{t}_{(i-1,j)} + Smell^{t}_{(i,j-1)} + Smell^{t}_{(i,j+1)} + \\ + Smell^{t}_{(i+1,j)} - 4 \times s_4 \times Smell^{t}_{(i,j)} \end{aligned} \tag{4}$$

and *RMN* expresses the remaining cells for the *Moore* neighbourhood to be created and is defined as:

$$\begin{aligned} RMN = Smell^{t}_{(i-1,j-1)} + Smell^{t}_{(i-1,j+1)} + Smell^{t}_{(i+1,j-1)} + \\ + Smell^{t}_{(i+1,j+1)} - 4 \times s_4 \times Smell^{t}_{(i,j)}, \end{aligned} \tag{5}$$

where: $s_2 = s_1 \times s_3$. The higher the values of parameters s_1, s_2, s_3 are, the quicker the diffusion of the food smell is. Parameter s_4 indicates the power of attraction of the Physarum by the food smell. The values of the $Smell^{t}_{(i,j)}$ variable range from 0 to N_{vs}, where $N_{vs} \in N$ is a determine integer number. In a cell $c_{i,j}$, the greater the value N_{vs} is, the more food smell exists in this particular cell.

Figure 3a illustrates a space, more precisely a corridor. The black color represents the walls, the blue the exits and the yellow face the human. In Figure 3 the food diffusion of this space is demonstrated. The white color represents the food source, while the more shaded a cell is, the less food smell exists in that cell ($N_{ns} \to 0$). All colors expect the red one represent the active cells K. Red color represents the walls of the environment, where no food smell has diffused.

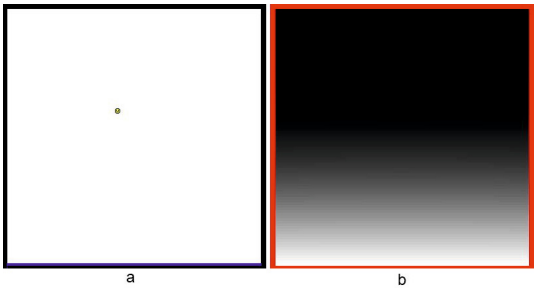


Fig. 3 Space before and after food diffusion: a) A corridor –white color– with a human –yellow face– with multiple exits which are represented with blue color. b) Food diffusion: The white color represents the food sources, whereas the more the color verges the black, the less the food smell has spread. Red color represents the walls, where no food smell has diffused

Organism's Growth. After the food diffusion, the growth of the microorganism takes place. It is important for the smell of the food sources to be spread primarily, since the growth of each plasmodium depends on the food diffusion. That means that each cytoplasm is directed to the closest food source in order to be fed. Thus, the microorganism should be able to detect the direction the food comes from and

given that the *P.* has only the ability to detect the smell of the food, the food smell should have been diffused.

Once more, the diffusion is carried out only in the active cells:

$$A = \left\{ c_{i,j}^{A_1}, c_{i,j}^{A_2}, c_{i,j}^{A_3}, \ldots, c_{i,j}^{A_K} \right\},$$

meaning that each microorganism can spread at whichever cell is not occupied by an obstacle or a wall. However, the organism can be propagated in a cell that is occupied by another organism, namely by another human. This happens, since the organisms during the diffusion process are independent from each other and they are considered as separate organisms, not able to merge. This could be feasible, considering as many nutrient-free substrates as they number of humans/ microorganisms.

- Let $H = \{H_1, H_2, , H_3, \ldots\}$ be the humans of the space Ω.
- Let N_H be the number of H and therefore of the organism's, so that: $H = \{H_1, H_2, H_3, \ldots, H_{N_H}\}$.
- Let $\left\{ c_{i,j}^{H_1}, c_{i,j}^{H_2}, c_{i,j}^{H_3}, \ldots, c_{i,j}^{H_{N_H}} \right\}$ be the cells, where the humans H are located.
- Define $ConOrg_{(i,j)}$ as the concentration of the plasmodium of Physarum at each position $\left\{ c_{i,j}^{H_1}, c_{i,j}^{H_2}, c_{i,j}^{H_3}, \ldots, c_{i,j}^{H_{N_H}} \right\}$ that the H human is located.

The growth of the micro–organism starts from the corresponding position $c_{i,j}^{H_\tau}$ of the human $H_\tau, \tau \in [1, \ldots, N_H]$ and continues/ proceeds for all cells of the environment according to:

$$ConOrg_{(i,j)}^{t+1} = ConOrg_{(i,j)}^{t} + c_1 \times MVN + c_2 \times LMN, \tag{6}$$

where MVN expresses the *von Neumann* neighbourhood and it is defined as follows:

$$\begin{aligned} MVN = &\left(1 + DF_{(i-1,j)}^{t}\right) \times ConOrg_{(i-1,j)}^{t} + \left(1 + DF_{(i,j-1)}^{t}\right) \times ConOrg_{(i,j-1)}^{t} + \\ &+ \left(1 + DF_{(i,j+1)}^{t}\right) \times ConOrg_{(i,j+1)}^{t} + \left(1 + DF_{(i+1,j)}^{t}\right) \times ConOrg_{(i+1,j)}^{t} - \\ &-4 \times c_4 \times ConOrg_{(i,j)}^{t} \end{aligned} \tag{7}$$

and LMV expresses the cells left for the *Moore* neighbourhood to be completed and is defined as follows:

$$\begin{aligned} LMV = &\left(1 + DF_{(i-1,j-1)}^{t}\right) \times ConOrg_{(i-1,j-1)}^{t} + \left(1 + DF_{(i-1,j+1)}^{t}\right) \times ConOrg_{(i-1,j+1)}^{t} + \\ &+ \left(1 + DF_{(i+1,j-1)}^{t}\right) \times ConOrg_{(i+1,j-1)}^{t} + \left(1 + DF_{(i+1,j+1)}^{t}\right) \times ConOrg_{(i+1,j+1)}^{t} - \\ &-4 \times c_4 \times ConOrg_{(i,j)}^{t}, \end{aligned} \tag{8}$$

where $DF_{(i,j)}^{t}, \forall (i,j) \in F \subseteq \Omega$ denotes the direction of the food smell $Smell_{(i,j)}^{t}$ that the organism H_τ detects for each cell at each time step t. The higher the values of parameters c_1, c_2, c_3 are, the quicker the growth of the organism's mass, respectively,

is. The c_4 parameter indicates the power of attraction of the Physarum by the food smell.

If there is no food smell then

$$DF^t_{(i,j)} = 0, \forall i,j \in [-1,1] : i \; or \; j \neq 0 \tag{9}$$

and if the food smell comes from the direction: (k,l), where $k,l \in [-1,1]$ then

$$DF^t_{(i+k,j+l)} = c_5 \; and \; DF^t_{(i-k,j-l)} = -c_5, \forall i,j \in [-1,1] : i \; or \; j \neq 0 \tag{10}$$

so that:

$$\sum_{\substack{i,j=-1 \\ i||j\neq 0}}^{1} DF^t_{i,j} = 0, \tag{11}$$

where it holds that: $c_2 = c_1 \times c_3$ and the variable c_5 indicates the influence the food smell $Smell^t_{(i,j)}$ has to the organism's growth $ConOrg^t_{(i,j)}$. The values of $ConOrg^t_{(i,j)}$ range from 0 to N_{vo}, where N_{vo} is a determine integer number. In a cell $c_{i,j}$, the greater the value N_{vo} is, the more mass of the organisms propagated in this particular cell.

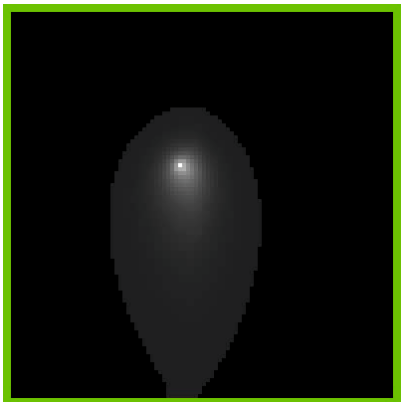


Fig. 4 The growth of the plasmodium *Physarum polycephalum*: The green color represents the walls, where no organism has been grown, whereas in the rest of the space, the more shadowed the color is, the less organism has propagated in that cell

Figure 4 illustrates the growth of the Physarum in the environment presented in Figure 3a. The white color represents the accumulated mass of the organism in its initial position (human's position), while the more shaded a cell is, the less the plasmodia have been grown in that cell ($N_{vo} \to 0$). All colors expect the green one represent the active cells A. Green color represents the walls of the environment, where there is no possibility for the organism to spread.

Protoplasmic Tubes' Creation. After the growth process, the creation of protoplasmic tubes follows:

- Let $C = \{C_1, C_2, C_3, \ldots\}$ be the set of the total protoplasmic tubes formed by all organisms.
- Let N_C denote the total number of tubes C. Hence the set C is defined as: $C = C_1, C_2, C_3, \ldots, C_{N_C}$
- Let $T_H \subseteq C$ denote the total tubes of the human H. $T_H = C_1, C_2, C_3, \ldots$.

Consider the Physarum/human $H_\tau, \tau \in [1, \ldots, N_H]$ located in the cell $C_{i,j}^{H_\tau}$ and the exitfood source $F_\lambda, \lambda \in [1, \ldots, s]$. It holds that:

$$\forall F_\lambda, H_\tau, \lambda \in [1, \ldots, s], \tau \in [1, \ldots, N_H] \exists_{=1} C_\lambda^{H_\tau}, \tag{12}$$

where $C_\lambda^{H_\tau}$ denotes the tube of the human H_τ with the exit F_λ.

- Let $T_H \subseteq C$ denote the total tubes of the human H. It holds that:

$$T_H = \left\{C_{F_1}^H, C_{F_2}^H, \ldots, C_{F_s}^H\right\}, \text{ where } s \to N_{umber} \text{ of } F. \tag{13}$$

Therefore:

$$C = \bigcap_{\tau=1}^{N_H} T_{H_\tau} = \left\{C_{F_1}^{H_1}, C_{F_2}^{H_1}, \ldots, C_{F_s}^{H_1}, C_{F_1}^{H_2}, C_{F_2}^{H_2}, \ldots, C_{F_s}^{H_2}, \ldots, C_{F_1}^{H_{N_H}}, C_{F_2}^{H_{N_H}}, \ldots, C_{F_s}^{H_{N_H}}\right\}, \tag{14}$$

where $\tau = 1, \ldots, N_H$ and $s \to N_{umber}$ of F,
therefore, the set C consists of all tubes created for all humans for all exits. Hence, the number N_C of elements of the set C is defined as:

$$N_C = s \times N_H. \tag{15}$$

That means that each organism –human– H creates as many tubes T as the number of exits s of the environment Ω.

Consider the Physarum/human $H_\tau, \tau \in [1, \ldots, N]$, located in the cell $C_{i,j}^{H_\tau}$ and the exit/food source $F_\lambda, \lambda \in [1, \ldots, s]$ located in the cell $C_{i,j}^{F_\lambda}$. The direction $D_{i,j}^t, \forall (i,j) \in A \subseteq \Omega$ of the food smell $Smell_{i,j}^t$ as described in Equations 8–11 resulted in the proper diffusion of the organism towards the exits F. The values, the diffusion $ConOrg_{i,j}$ of the H_τ organism –so called $ConOrg_{i,j}^{H_\tau}$ –(according to Equation 3) received, represent the concentration of the organism in each cell $c_{i,j}$.

Each tube $C_{F_\lambda}^{H_\tau}$, $\lambda \in [1, \ldots, s]$, $\tau \in [1, \ldots, N_H]$ consists of successional cells as follows:

$$C_{F_\lambda}^{H_\tau} = C_{i,j}^{H_\tau}, \ldots, C_{i,j}^{F_\lambda} \ \lambda \in [1, \ldots, s], \ \tau \in [1, \ldots, N_H]. \tag{16}$$

In particular, the construction of the protoplasmic tube $C_{F_\lambda}^{H_\tau}$ starts from the cell $C_{i,j}^{H_\tau}$, where the human H_τ and hence the main mass of the organism is located. Then, of all the adjacent cells of the initial cell $C_{i,j}^{H_\tau}$, it is chosen the cell $C_{i_{new}, j_{new}}^{H_\tau}$ that:

$$ConOrg_{i_{new},j_{new}} = max\left\{ConOrg_{i+k,j+n}\right\} : c_{i+k,j+n} \notin C_{F_\lambda}^{H_\tau}, \tag{17}$$

where $k,n \in [-1,1] : k||n \neq 0$.

That means that from all adjacent cells, the cell chosen $C_{i_{new},j_{new}}^{H_\tau}$ is the one, in which the mass of the organism is the maximum as long as it does not already belong the tube cells, namely as long as the tube does not contain this cell. When this cell is chosen, the same process is followed in order for the tube $C_{F_\lambda}^{H_\tau}$ to be created. The construction ends to the F_λ food source, namely to the cell $C_{i,j}^{F_\lambda}$.

The same procedure is followed for the specific human H_τ for all the food sources\exits F_s. Thus, the set $T_{H_\tau} = \left\{C_{F_1}^{H_\tau}, C_{F_2}^{H_\tau}, \ldots, C_{F_s}^{H_\tau}\right\}$ with all the possible paths for the human to follow is created.

Figure 5 demonstrates the tubes created. Figure 5a illustrates a real experiment, where a microorganism is located in a nutrient–poor environment and creates a protoplasmic tube that connects its location\position with the food source located to the other side of the substrate. Figure 5b presents a space with one human and one exit. The human, which corresponds to the microorganism, is drawn with a yellow face, the walls and obstacles with black color and the tube with green.

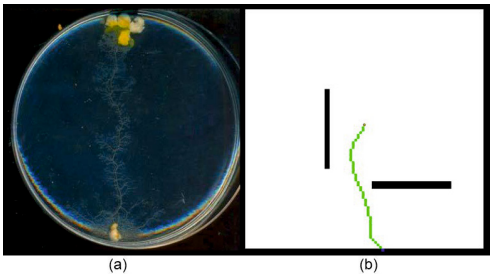


Fig. 5 Creation of tubes: a) An example of plasmodium attracted by source of nutrients [2]. Initially the plasmodium is placed in the southern part of disk and a group of intact oat flakes in the northern part. Plasmodium propagates towards the intact flakes and occupies them b) The plasmodium creates a tube with the exit demonstrated by green color

Figure 6 demonstrates the tubes created in cases where more than one human or more than one exits exist in the environment. Figure 6a illustrates a space with one human and two exits, while in Figure 6b there are two humans and two exits in the environment. In both cases the walls and the obstacles of the space are drawn with black color, the human with a yellow face and the tubes with green color. Consequently, it is quite obvious that each human, who imitates the functionality of the microorganism, is connected with protoplasmic tubes with every exit of the space.

However, only one tube is able to indicate the path the human will follow, given that a human follows one route. Hence, the most appropriate and representative of each humans motion tube should be selected.

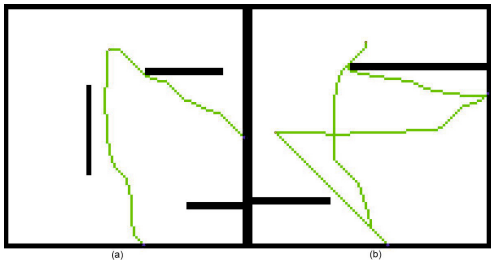


Fig. 6 Tube construction: a) A human in an environment with two exits. For each exit an individual tube is created b) Two humans with the corresponding tubes towards the exits

Selection of Shortest Tube. As it is made clear, the number of elements of each tube $C_{F_\lambda}^{H_\tau}, \forall\lambda \in [1,\ldots,s]$, $\forall\tau \in [1,\ldots,N_H]$ of the set $T_{H_\tau} = C_{F_1}^{H_\tau}, C_{F_2}^{H_\tau}, \ldots, C_{F_s}^{H_\tau}$ for a human H_τ is not a static number, but it depends on the number of cells $C_{i,j}$ the tube $C_{F_\lambda}^{H_\tau}$ consists of.

- Let assume $NCel_{C_{F_\lambda}^{H_\tau}}^{H_\tau}$ denote the number of cells of each tube for each human H_τ. It holds that:

$$1 \leq NCel_{C_{F_\lambda}^{H_\tau}}^{H_\tau} \leq K \subset N_{C_{i,j}}. \tag{18}$$

Thus, for the chosen tube $C_{F_\lambda ch}^{H_\tau}$ of a human H_τ it holds that:

$$NCel_{C_{F_\lambda}^{H_\tau}}^{H_\tau} = min\left\{NCel_{C_{F_1}^{H_\tau}}^{H_\tau}, NCel_{C_{F_2}^{H_\tau}}^{H_\tau}, \ldots, NCel_{C_{F_s}^{H_\tau}}^{H_\tau}\right\}, \tag{19}$$

which can be explained as follows: The chosen tube $C_{F_\lambda ch}^{H_\tau}$ for a human H_τ is the one that contains the minimum number of elements $\left\{c_{i,j}^{H_\tau}, \ldots, c_{i,j}^{F_\lambda}\right\}$, i.e. the minimum number of cells, namely the tube that corresponds to the shortest path. In case two or more tubes consist of equal number of cells, then the one with the smallest Euclidean distance is chosen. Figure 7 demonstrates such an example. In an environment with two humans and two exits each human chooses as target exit the one that is located nearest to him.

Move towards Tube. The next step of the evacuation algorithm contains the human's movement along the path/tube created and selected earlier. The movement process depends on whether or not the position the human is about to go is occupied. Therefore, at each time step t each human checks whether or not she/he is able to move. If it is so, the human follows the path to the exit. If not, then the she/he has two options:

1. To remain motionless. This option connotes that the human waits until the position she/he wants to move into will be released.
2. To follow a new path that leads to the closest exit.

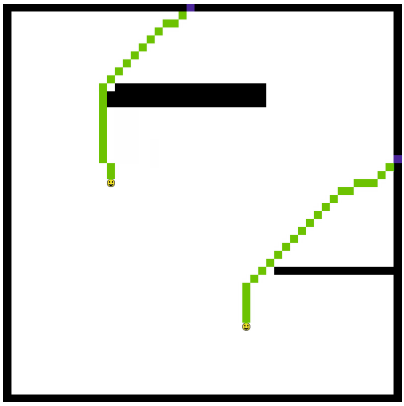


Fig. 7 An environment with two separate exits. From all tubes created for each human, only the one that connects him with the closest exit is chosen

To sum up, the proposed bio–inspired evacuation model is composed by: the food diffusion, the organism's growth, the creation of tubes for each organism, the selection of optimum tube for each human and finally by the movement of all humans at each time step. These steps are repeated as many times needed until all humans exit the environment.

4 Simulation Results for the Proposed Evacuation Model

As it is already known (Weidmann, 1993) the human velocity in areas with low people density approximately is set to be: $V_H = 1.34$ ms^{-1}. What is more, as already mentioned, each human occupies a single cell of the space that according to [11] corresponds to 0.4×0.4m^2. Given that at one time step of the method each human is considered to be able to move by one cell, it can be concluded that each time step of the method approximately corresponds to 0.3 s –specifically to 0.298507 s in a real time experiment. The carried out simulation cases were enough and differentiate significantly so as to examine different well known test cases in corwd evacuation modeling. In specific five different representative examples are demonstrated, through which the whole evacuation procedure can be easily understood as well as the results of the evacuation scenarios can be visualized. In all five examples listed below, the values of the pre-mentioned parameters are set to be (see Table 1):

Table 1 General parameters of all evacuation scenarios regarding diffusion of the food as well as of the plasmodium of *Physarum*

Parameters of Food Diffusion	Values	Parameters of Organism's Diffusion	Values
s_1	0.1	c_1	0.1
s_3	0.01	c_3	0.01
$s_2 = s_1 \times s_3$	0.001	$c_2 = c_1 \times c_3$	0.001
s_4	1	c_4	1
–	–	c_5	0.3

4.1 Experiment 1

Initially, the environment to be evacuated was created. As first example a simple space is selected. The space is actually a corridor with multiple exits and without any obstacles. The parameters of this environment are shown in Table 2 below, while the images of the experiment are shown in Figure 8. As it can be easily observed all humans manage to exit the space successfully.

Table 2 The parameters of evacuation scenario No. 1

Parameters	Case 1
$L_1(width)$	$21.6m$
$L_2(height)$	$40m$
N_H	192
$s(exits)$	50
$N_{1C_{i,j}} = \frac{L_1}{0.4}$	54 cells
$N_{2C_{i,j}} = \frac{L_2}{0.4}$	100 cells
$N_{C_{i,j}} = N_{1C_{i,j}} \times N_{2C_{i,j}}$	54×100 cells
$Obstacles$	$None$
$t(Time\ Steps)$	88
$Real\ time$	$26.26s$

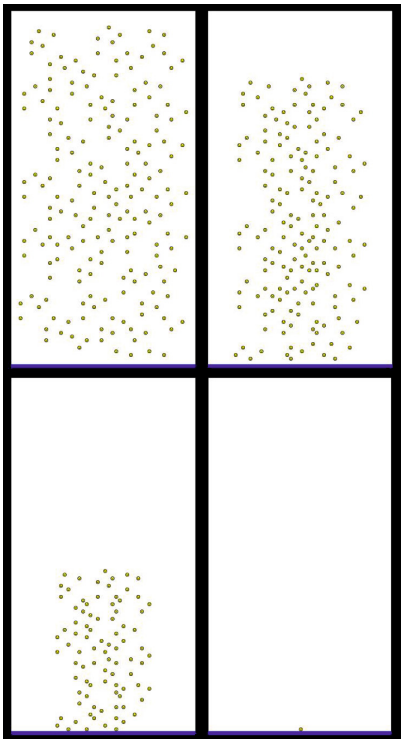


Fig. 8 Evacuation Scenario No. 1: A corridor with 50 exits, 192 humans and no obstacles

4.2 Experiment 2

In the second evacuation scenario a corridor is once more selected. However, this time the corridor consists of only one exit, meaning that at each time step only one human is able to exit the environment. The parameters of this environment are shown in Table 3 below, while the images of the experiment are shown in Figure 9. The main goal of this experiment is to point out the congestion in the exit. In such cases effects of building up pressure are observed (Helbing, Farkas, & Vicsek, Simulating Dynamical Features of Escape Panic, 2000), since humans are pressed to each other. What is more, in this case the prominent arching effects are observed as in real experiments.

4.3 Experiment 3

In this evacuation scenario an environment with two (2) exits is selected. These exits are not next to each other, they are placed on different sides of the under study

Table 3 The parameters of evacuation scenario No. 2

Parameters	Values
$L_1(width)$	$21.6m$
$L_2(height)$	$40m$
N_H	192
$s(exits)$	1
$N_{1C_{i,j}} = \frac{L_1}{0.4}$	54 cells
$N_{2C_{i,j}} = \frac{L_2}{0.4}$	100 cells
$N_{C_{i,j}} = N_{1C_{i,j}} \times N_{2C_{i,j}}$	54×100 cells
Obstacles	*None*
$t(Time\ Steps)$	192
Real time	$57.3s$

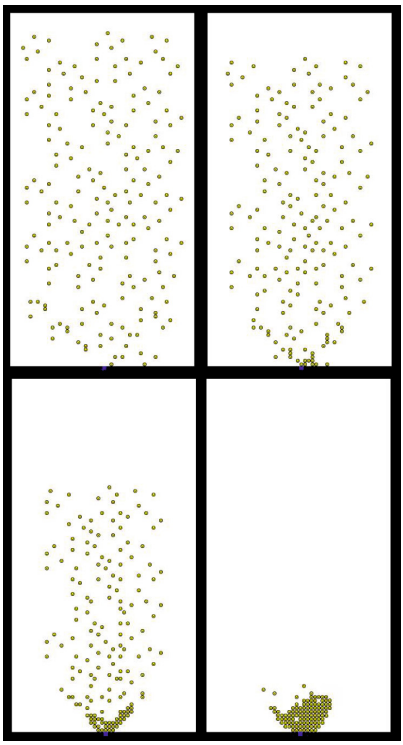


Fig. 9 Evacuation Scenario 2: A corridor with only one (1) exit, 192 humans and no obstacles

space in order to clarify the fact that each human is able to choose the most appropriate exit. In that way, the so called clustering effect is observed (Helbing, Farkas, & Vicsek, Simulating Dynamical Features of Escape Panic, 2000). This is happening since humans are divided into groups: humans that select to exit from the first exit and other that leave the building from the other exit. The parameters of this environment are shown in Table 4, while snapshots of the experiment are shown in Figure 10.

Table 4 The parameters of evacuation scenario No. 3

Parameters	Values
$L_1(width)$	$20.0m$
$L_2(height)$	$20.0m$
N_H	99
$s(exits)$	2
$N_{1C_{i,j}} = \frac{L_1}{0.4}$	50 cells
$N_{2C_{i,j}} = \frac{L_2}{0.4}$	50 cells
$N_{C_{i,j}} = N_{1C_{i,j}} \times N_{2C_{i,j}}$	50×50 cells
$Obstacles$	Yes
$t(Time\ Steps)$	59
$Real\ time$	$17.62s$

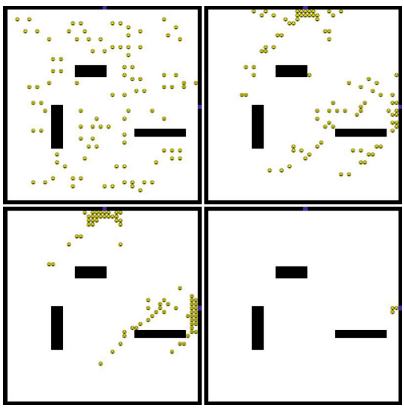


Fig. 10 Evacuation Scenario No. 3: An environment with 2 exits, 99 humans and some randomly placed obstacles

4.4 Experiment 4

In this evacuation scenario an environment with U shaped obstacles and three (3) adjacent located in the north side of the under study space exits is examined. The simulation of a scenario with U shaped obstacles is quite crucial, since humans may need to go in opposite direction of the exit in order to avoid the U-shaped obstacle. The parameters of this environment are shown in Table 5, while the images of the experiment are shown in Figure 11. As it can be observed, in this case scenario clogging effects at bottlenecks appear (Helbing, Farkas, & Vicsek, Simulating Dynamical Features of Escape Panic, 2000). However, the model is able to overcome such difficulties, given that humans evacuate successfully the place.

Table 5 The parameters of evacuation scenario No. 4

Parameters	Values
$L_1(width)$	$20.0m$
$L_2(height)$	$20.0m$
N_H	161
$s(exits)$	3
$N_{1C_{i,j}} = \frac{L_1}{0.4}$	50 cells
$N_{2C_{i,j}} = \frac{L_2}{0.4}$	50 cells
$N_{C_{i,j}} = N_{1C_{i,j}} \times N_{2C_{i,j}}$	50×50 cells
$Obstacles$	$U--shaped$
$t(Time\ Steps)$	69
$Real\ time$	$20.5s$

4.5 Experiment 5

In this evacuation scenario the space consists of multiple rooms. The rooms are the one inside the other and therefore humans should leave all rooms before reaching the exit of the space. The parameters of this environment are shown in Table 6 below, while the images of the experiment are shown in Figure 14. As it can be observed, in this case scenario clogging and arching effects appear (Helbing, Farkas, & Vicsek, Simulating Dynamical Features of Escape Panic, 2000). The simulation results demonstrated in the following Figure 12 indicate the ability of the model to respond properly to such scenarios.

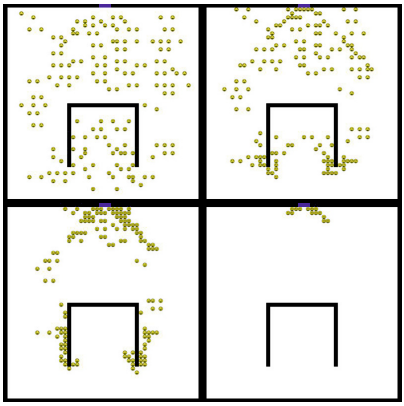


Fig. 11 Evacuation Scenario No. 4: A space with 3 consecutive exits and with a U - shaped obstacle

Table 6 The parameters of evacuation scenario No. 5

Parameters	Values
$L_1(width)$	$20.0m$
$L_2(height)$	$20.0m$
N_H	66
$s(exits)$	1
$N_{1C_{i,j}} = \frac{L_1}{0.4}$	50 cells
$N_{2C_{i,j}} = \frac{L_2}{0.4}$	50 cells
$N_{C_{i,j}} = N_{1C_{i,j}} \times N_{2C_{i,j}}$	50×50 cells
$Obstacles$	$walls$
$t(Time\ Steps)$	71
$Real\ time$	$21.2s$

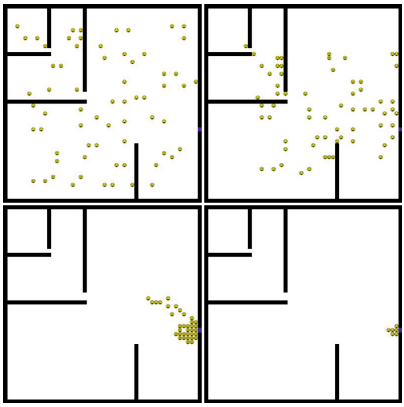


Fig. 12 Evacuation Scenario No. 5: A space with internal rooms and 3 consecutive corresponding exits

5 Conclusions

In this chapter a bio–inspired crowd evacuation method based on CA and bio–mimicking Physarum's behaviour was described. The presented model take into account while mimicking the Physarum foraging process, the food diffusion, the organism's growth, the creation of tubes for each organism, the selection of optimum tube for each human in correspondence to the under study crowd evacuation and finally the movement of all humans at each time step towards near exit. The proposed evacuation model combines microscopic and macroscopic characteristics enabling the successful simulation of crowd evacuation taking into account key feature of evacuation of the places crowded by humans, like unity in diversity. In specific, several simulation scenarios were examined both in virtual indoor environments. The corresponding results have proven in a qualitative way the efficiency of the proposed model in terms of producing random to coherent motion of individuals due to common purpose, collective effects and blockings near exits in all the examined cases. Moreover, different behaviours during overcrowdness at one exit, as well as due to fire spreading, were successfully reproduced in some of the provided experiments.

References

1. Adamatzky, A.: Physarum machine: Implementation of a kolmogorov-uspensky machine on a biological substrate. Parallel Processing Letters 17(4), 455–467 (2007)
2. Adamatzky, A.: Physarum machines: computers from slime mould, vol. 74. World Scientific, Singapore (2010)
3. Adamatzky, A.: Slime mold solves maze in one pass, assisted by gradient of chemo-attractants. IEEE Transactions on NanoBioscience 11(2), 131–134 (2012)

4. Adamatzky, A.: Route 20, autobahn 7 and physarum polycephalum: Approximating longest roads in usa and germany with slime mould on 3d terrains. arXiv preprint arXiv:1211.0519. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics (2013) (in press)
5. Adamatzky, A., Jones, J.: Road planning with slime mould: if physarum built motorways it would route m6/m74 through newcastle. I. J. Bifurcation and Chaos 20(10), 3065–3084 (2010)
6. Adamatzky, A., Schumann, A.: Physarum spatial logic. New Mathematics and Natural Computation 07(03), 483–498 (2011)
7. Aubé, F., Shield, R.: Modeling the effect of leadership on crowd flow dynamics. In: Sloot, P.M.A., Chopard, B., Hoekstra, A.G. (eds.) ACRI 2004. LNCS, vol. 3305, pp. 601–611. Springer, Heidelberg (2004)
8. Bandini, S., Manzoni, S., Vizzari, G.: Situated cellular agents: A model to simulate crowding dynamics. IEICE Transactions on Information and Systems 87(3), 669–676 (2004)
9. Braun, A., Musse, S.R., de Oliveira, L.P.L., Bodmann, B.E.: Modeling individual behaviors in crowd simulation. In: 16th International Conference on Computer Animation and Social Agents, pp. 143–148. IEEE (2003)
10. Brogan, D.C., Hodgins, J.K.: Simulation level of detail for multiagent control. In: Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1, pp. 199–206. ACM (2002)
11. Burstedde, C., Klauck, K., Schadschneider, A., Zittartz, J.: Simulation of pedestrian dynamics using a two-dimensional cellular automaton. Physica A: Statistical Mechanics and its Applications 295(3), 507–525 (2001)
12. Chenney, S.: Flow tiles. In: Proceedings of the 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 233–242. Eurographics Association (2004)
13. Chopard, B., Droz, M.: Cellular automata modeling of physical systems, vol. 122. Springer (1998)
14. Daoliang, Z., Lizhong, Y., Jian, L.: Exit dynamics of occupant evacuation in an emergency. Physica A: Statistical Mechanics and its Applications 363(2), 501–511 (2006)
15. Feynman, R.P.: Simulating physics with computers. International Journal of Theoretical Physics 21(6), 467–488 (1982)
16. Georgoudas, I., Sirakoulis, G.C., Scordilis, E., Andreadis, I.: A cellular automaton simulation tool for modelling seismicity in the region of xanthi. Environmental Modelling & Software 22(10), 1455–1464 (2007)
17. Georgoudas, I.G., Kyriakos, P., Sirakoulis, G.C., Andreadis, I.T.: An fpga implemented cellular automaton crowd evacuation model inspired by the electrostatic-induced potential fields. Microprocessors and Microsystems 34(7), 285–300 (2010)
18. Georgoudas, I.G., Sirakoulis, G.C., Andreadis, I.T.: A simulation tool for modelling pedestrian dynamics during evacuation of large areas. In: Maglogiannis, I., Karpouzis, K., Bramer, M. (eds.) Artificial Intelligence Applications and Innovations. IFIP, vol. 204, pp. 618–626. Springer, Heidelberg (2006)
19. Gunji, Y.P., Shirakawa, T., Niizato, T., Haruna, T.: Minimal model of a cell connecting amoebic motion and adaptive transport networks. Journal of Theoretical Biology 253(4), 659–667 (2008)
20. Helbing, D., Farkas, I., Vicsek, T.: Simulating dynamical features of escape panic. Nature 407(6803), 487–490 (2000)
21. Henderson, L.: The statistics of crowd fluids. Nature 229, 381–383 (1971)

22. Henein, C.M., White, T.: Agent-based modelling of forces in crowds. In: Davidsson, P., Logan, B., Takadama, K. (eds.) MABS 2004. LNCS (LNAI), vol. 3415, pp. 173–184. Springer, Heidelberg (2005)
23. Hoogendoorn, S.P.: Pedestrian travel behavior modeling. In: 10th International Conference on Travel Behavior Research, Lucerne (2003)
24. Jendrsczok, J., Ediger, P., Hoffmann, R.: A scalable configurable architecture for the massively parallel gca model. International Journal of Parallel, Emergent and Distributed Systems 24(4), 275–291 (2009)
25. Jian, L., Lizhong, Y., Daoliang, Z.: Simulation of bi-direction pedestrian movement in corridor. Physica A: Statistical Mechanics and its Applications 354, 619–628 (2005)
26. Jones, J.: Approximating the behaviours of *physarum polycephalum* for the construction and minimisation of synthetic transport networks. In: Calude, C.S., Costa, J.F., Dershowitz, N., Freire, E., Rozenberg, G. (eds.) UC 2009. LNCS, vol. 5715, pp. 191–208. Springer, Heidelberg (2009)
27. Kalogeiton, V.S., Papadopoulos, D.P., Sirakoulis, G.C.: Hey physarum! can you perform slam? IJUC 10(4), 271–293 (2014)
28. Karafyllidis, I.: A model for the prediction of oil slick movement and spreading using cellular automata. Environment International **23**(6), 839 – 850 (1997). doi:http://dx.doi.org/10.1016/S0160-41209700096-2
29. Karafyllidis, I., Thanailakis, A.: A model for predicting forest fire spreading using cellular automata. Ecological Modelling 99(1), 87–97 (1997)
30. Karafyllidis, I., Thanailakis, A.: A model for predicting forest fire spreading using cellular automata. Ecological Modelling 99(1), 87–97 (1997), doi:http://dx.doi.org/10.1016/S0304-38009601942-4
31. Kirchner, A., Nishinari, K., Schadschneider, A.: Friction effects and clogging in a cellular automaton model for pedestrian dynamics. Physical Review E **67**(5) 056, 122 (2003)
32. Lindzey, G.E., Aronson, E.E. (eds.): The handbook of social psychology. Addison-Wesley (1968)
33. Liu, Y., Zhang, Z., Gao, C., Wu, Y., Qian, T.: A *physarum* network evolution model based on IBTM. In: Tan, Y., Shi, Y., Mo, H. (eds.) ICSI 2013, Part II. LNCS, vol. 7929, pp. 19–26. Springer, Heidelberg (2013)
34. Mardiris, V., Sirakoulis, G.C., Mizas, C., Karafyllidis, I., Thanailakis, A.: A cad system for modeling and simulation of computer networks using cellular automata. IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews 38(2), 253–264 (2008)
35. Milazzo, J.S., Rouphail, N.M., Hummer, J.E., Allen, D.P.: Effect of pedestrians on capacity of signalized intersections. Transportation Research Record: Journal of the Transportation Research Board 1646(1), 37–46 (1998)
36. Musse, S.R., Thalmann, D.: Hierarchical model for real time simulation of virtual human crowds. IEEE Transactions on Visualization and Computer Graphics 7(2), 152–164 (2001)
37. Nakagaki, T., Yamada, H., Tóth, A.: Intelligence: Maze-solving by an amoeboid organism. Nature 407(6803), 470 (2000)
38. Nakagaki, T., Yamada, H., Toth, A.: Path finding by tube morphogenesis in an amoeboid organism. Biophysical Chemistry 92(1), 47–52 (2001)
39. Nakagaki, T., Yamada, H., Ueda, T.: Interaction between cell shape and contraction pattern in the i physarum plasmodium. Biophysical Chemistry 84(3), 195–204 (2000)
40. Nishinari, K., Sugawara, K., Kazama, T., Schadschneider, A., Chowdhury, D.: Modelling of self-driven particles: Foraging ants and pedestrians. Physica A: Statistical Mechanics and its Applications 372(1), 132–141 (2006)

41. Paris, S., Donikian, S.: Activity-driven populace: a cognitive approach to crowd simulation. IEEE Computer Graphics and Applications 29(4), 34–43 (2009)
42. Perez, G.J., Tapang, G., Lim, M., Saloma, C.: Streaming, disruptive interference and power-law behavior in the exit dynamics of confined pedestrians. Physica A: Statistical Mechanics and its Applications 312(3), 609–618 (2002)
43. Schultz, M., Lehmann, S., Fricke, H.: A discrete microscopic model for pedestrian dynamics to manage emergency situations in airport terminals. In: Pedestrian and Evacuation Dynamics 2005, pp. 369–375. Springer (2007)
44. Schumann, A., Adamatzky, A.: Toward semantical model of reaction-diffusion computing. Kybernetes 38(9), 1518–1531 (2009)
45. Shao, W., Terzopoulos, D.: Autonomous pedestrians. Graphical Models 69(5), 246–274 (2007)
46. Shirakawa, T., Adamatzky, A., Gunji, Y.P., Miyake, Y.: On simultaneous construction of voronoi diagram and delaunay triangulation by physarum polycephalum. International Journal of Bifurcation and Chaos 19(09), 3109–3117 (2009)
47. Shirakawa, T., Adamatzky, A., Gunji, Y.P., Miyake, Y.: On simultaneous construction of voronoi diagram and delaunay triangulation by physarum polycephalum. I. J. Bifurcation and Chaos 19(9), 3109–3117 (2009)
48. Sirakoulis, G.C.: A tcad system for vlsi implementation of the cvd process using vhdl. Integration, the VLSI Journal 37(1), 63–81 (2004)
49. Sirakoulis, G.C., Bandini, S. (eds.): ACRI 2012. LNCS, vol. 7495. Springer, Heidelberg (2012)
50. Sirakoulis, G.C., Karafyllidis, I., Thanailakis, A.: A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. Ecological Modelling 133(3), 209–223 (2000)
51. Sirakoulis, G.C., Karafyllidis, I., Thanailakis, A.: A cad system for the construction and vlsi implementation of cellular automata algorithms using vhdl. Microprocessors and Microsystems 27(8), 381–396 (2003)
52. Spezzano, G., Talia, D., Di Gregorio, S., Rongo, R., Spataro, W.: A parallel cellular tool for interactive modeling and simulation. IEEE Computational Science & Engineering 3(3), 33–43 (1996)
53. Stephenson, S.L., Stempen, H., Hall, I.: Myxomycetes: a handbook of slime molds. Timber press Portland, Oregon (1994)
54. Tero, A., Kobayashi, R., Nakagaki, T.: A mathematical model for adaptive transport network in path finding by true slime mold. Journal of Theoretical Biology 244(4), 553 (2007)
55. Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D.P., Fricker, M.D., Yumiki, K., Kobayashi, R., Nakagaki, T.: Rules for biologically inspired adaptive network design. Science 327(5964), 439–442 (2010)
56. Toffoli, T.: Cam: A high-performance cellular-automaton machine. Physica D: Nonlinear Phenomena 10(1), 195–204 (1984)
57. Tsompanas, M.A.I., Sirakoulis, G.C.: Modeling and hardware implementation of an amoeba-like cellular automaton. Bioinspiration & Biomimetics 7(3), 036,013 (2012)
58. Tsuda, S., Aono, M., Gunji, Y.P.: Robust and emergent physarum logical-computing. Biosystems 73(1), 45–55 (2004)
59. Varas, A., Cornejo, M., Mainemer, D., Toledo, B., Rogan, J., Munoz, V., Valdivia, J.: Cellular automaton model for evacuation process with obstacles. Physica A: Statistical Mechanics and its Applications 382(2), 631–642 (2007)
60. Vichniac, G.Y.: Simulating physics with cellular automata. Physica D: Nonlinear Phenomena 10(1), 96–116 (1984)

61. Vizzari, G., Manenti, L., Crociani, L.: Adaptive pedestrian behaviour for the preservation of group cohesion. Complex Adaptive Systems Modeling 1(1), 1–29 (2013)
62. Von Neumann, J., Burks, A.W., et al.: Theory of self-reproducing automata. University of Illinois press Urbana (1966)
63. Weifeng, F., Lizhong, Y., Weicheng, F.: Simulation of bi-direction pedestrian movement using a cellular automata model. Physica A: Statistical Mechanics and its Applications 321(3), 633–640 (2003)
64. Wilding, N.B., Trew, A., Hawick, K., Pawley, G.: Scientific modeling with massively parallel simd computers. Proceedings of the IEEE 79(4), 574–585 (1991)
65. Wolfram, S.: Theory and applications of cellular automata. Advanced Series on Complex Systems. World Scientific Publication, Singapore (1986)
66. Yang, L., Zhao, D., Li, J., Fang, T.: Simulation of the kin behavior in building occupant evacuation based on cellular automaton. Building and Environment 40(3), 411–415 (2005)
67. Yu, Y., Song, W.: Cellular automaton simulation of pedestrian counter flow considering the surrounding environment. Physical Review E 75(4), 046,112 (2007)
68. Yuan, W., Tan, K.H.: An evacuation model using cellular automata. Physica A: Statistical Mechanics and its Applications 384(2), 549–566 (2007)

Neural Net to Neuronal Network Memristor Interconnects

Ella Gale, Attya Iqbal, Jeffrey Davey, and Deborah Gater

Abstract. Hardware-based Artificial Intelligence (A.I.) has many potential applications in biomedical technology; for example, connecting expert systems to biosensors for real-time physiological monitoring of a range of biomarkers, or interfacing with the brain. We suggest that memristors are well-placed to interface directly with neurons to interconnect between computer hardware and the brain for 3 reasons: memristors are widely-touted as neuromorphic computing elements; memristor theory has been successfully used to model spike-time dependent plasticity and the Hodgkin-Huxley model of the neuron; and, the d.c. response of the memristor is a current spike. In this chapter we show that connecting a spiking memristor network to electrically active neuronal cells causes a change in the memristor network dynamics by: removing the memristor spikes, which we show is due to the effects of connection to aqueous medium; causing a change in current decay rate consistent with a change in memristor state; presenting more-linear $I-t$ dynamics; and increasing the memristor spiking rate, as a consequence of interaction with the active cells. This demonstrates that such cells are capable of communicating directly with memristors, without the need for computer translation.

Ella Gale
Bristol Robotics Laboratory, Bristol, UK
e-mail: ella.gale@brl.ac.uk

Jeffrey Davey · Attya Iqbal · Ella Gale
Centre for Research in Biosciences, University of the West of England, Bristol, UK

Deborah Gater
Department of Applied Mathematics and Sciences, Khalifa University of Science, Technology and Research, Abu Dhabi, UAE
e-mail: deborah.gater@kustar.ac.ae

K. Pancerz and E. Zaitseva (eds.), *Computational Intelligence, Medicine and Biology*,
Studies in Computational Intelligence 600, DOI: 10.1007/978-3-319-16844-9_8

1 Introduction

Artificial Intelligence, A.I., has been of use in medicine primarily through the adoption of expert systems, essentially giant look-up tables of symptoms, that aid in diagnosis [10, 36, 31, 34]. A neural net is a network of components, instantiated either via algorithms in software or physical hardware, that mimic the brain's operation and is thus of great interest in the development of advanced A.I.. A neuronal network is a network of living neurons (and often other neural cells, such as astrocytes) either within a living brain or as part of a simpler *in vitro* system of cultured neural cells. Previous attempts have demonstrated the feasibility of connecting an A.I. system directly to neuronal cells [38, 30]. In this chapter, we focus on a novel device, the memristor, that has possible uses both in the creation of hardware neural nets for A.I. and as the connection between a hardware neural net and a living neuronal cell network for uses in the treatment and monitoring of neurological disease.

The human body is partly electrical in nature, with nerve impulses being electrical signals, cardiac motion being controlled by electrical impulses and the brain 'computing' via electrical spikes (both current spikes across synapses and voltage spikes along axons). The development of systems that are capable of reliable, minimally invasive electrical interface with the brain and that can be incorporated into portable, wearable devices, would allow the creation of cognitive prosthetics to combat a variety of serious neurological conditions resulting from disease, ageing or injury. For example, an electrical 'circuit breaker' could reduce the symptoms of epilepsy and deep-brain stimulation has been shown to reduce epileptic symptoms [6], "prosthetic neurons" could replace neurons damaged by Alzheimer's Disease, and direct current stimulation has been shown to decrease the symptoms of Parkinson's disease [2].

Wearable expert systems connected to biosensors (devices capable of on-line detecting and reporting various physiological parameters, from blood pressure to glucose levels) could prove invaluable in monitoring numerous facets of health and disease. However, the monitoring of diseases with a neurological component may require a direct connection to the brain, either using a wearable EEG device or via another form of direct electrical contact. The second approach would enable the collection of finer data than is possible with an EEG, but can require drastic medical intervention, usually involving surgery to implant electrodes. However, it is possible that memristors will have some advantages over standard electrodes in this area: memristors can be manufactured at the nanoscale [35] to minimise disruption to delicate neural tissue, flexible [21] to reduce the mechanical strain between the cells and a rigid electrode, and manufactured from organic materials [22, 12] to reduce the potential for biofouling (where the body grows over metal implants reducing their effectiveness and lifespan).

1.1 The Memristor

The memristor [7] is a resistor that possesses a memory. The first physical memristor was reported in 2008 [35] and consisted of two layers of titanium dioxide: a stoichiometric (high resistance) semiconductor layer of TiO_2 and a doped sub-oxide (low resistance) layer of TiO_{2-x} (where x is $\tilde{3}\%$ of the available oxygen) sandwiched between platinum electrodes. Under an applied voltage the material can inter-convert between the two forms, $TiO_2 \longleftrightarrow TiO_{2-x}$, changing the relative amount of each. As the different oxides have different resistivities, the resistance of the device thus changes relative to the amount of charge driven by voltage it has been subjected to.

It is the change in resistance as a function of the charge passed through the device throughout its history that allows the memristor to have a memory[1]. In standard von Neumann computers, the processor and memory are separate hardware components. In contrast, in the brain the processing and memory happen in very close proximity. As the memristor possesses both memory and processing power[2] it has been suggested as a natural component to build neuromorphic (brain-like) processors [35]. It has even been suggested that the failure of computer science to build a computer on the same order of complexity as a brain in the same volume is the result of the lack of utilisation of the memristor. There are big projects to correct this such as DARPA's SyNAPSE program. The focus is on a hardware instantiation of the design as a hardware A.I. or neuromorphic computer would be faster and smaller than an equivalently powerful software A.I. system.

Most research has concentrated on using memristors to construct permanent memory for computer memory, however nanoscale processing reduces repeatability below that required for this type of storage. As the brain is capable of robust computation with imperfect components, a possible solution is to use neural-inspired engineering.

Previous work [15, 20, 16] has drawn attention to the short-term memory aspects of the memristor, which might be useful in copying neuronal networks. The d.c. response of the memristor is a current spike that contains the short-term memory of the memristor [15]; on a single memristor these spikes can be used to compute in a Boolean fashion [16], and networks of memristors exhibit brainwave-like dynamics and bursting spikes [20]. The similarity between memristor network dynamics, and brainwaves and neuronal spiking patterns suggests that the memristor could be a good choice for either copying the brain (i.e. creating a biomimetic neuromorphic computer) or interfacing with it.

[1] Theoretically the memory is a function of the entire history of the device [7], experimentally, it is the history since the device was zeroed.

[2] In that it can be used as part of a processor.

1.2 Action of Neuronal Networks

The brain is thought to work (in part) by forming and breaking connections between individual neurons. An individual neuron can have many connections (termed synapses) with other cells, and it is the synapse that controls the transmission or damping of electrical signals. The more a connection between neurons is used, the stronger the connection becomes and this process of learning is called Hebbian learning [23]. A process called spike time dependent plasticity [5] explains Hebbian learning. Each synapse is unidirectional: the neuron before the synapse is called the pre-synaptic neuron, the one after it is called the post-synaptic neuron. A synapse receives a signal in terms of a voltage spike with no knowledge of the origin of that signal. Thus, if a signal is received at the pre-synaptic neuron before the post-synaptic neuron, the synapse 'assumes' the signal at the post-synaptic neuron was caused by the signal at the pre-synaptic neuron and thus they are related. In response, it lowers the threshold for transmitting signals from the pre-synaptic neuron to the post-synaptic neuron. This is Hebbian learning, the oft-quoted concept that 'neurons that fire together, wire together', and accounts for the increased ease of recalling frequently referred-to information and improvements at frequently repeated tasks.

Physiologically, the synapse increases the number of synaptic clefts, which bud and release packets of chemical signalling agents that pass the signals across the synapse. However, as there are so many connections in the brain, it is possible that a signal can be received at the post-synaptic neuron from an origin other than the pre-synaptic neuron. Thus, if a signal is received outside of the period of time it would take for the chemical signals to cross the synapse, the synaptic weight does not change (this is why the process is spike-time dependent). If the signal is received at the post-synaptic neuron before the pre-synaptic neuron, the two neurons are not related and the synaptic weight is decreased[3]. The neuronal network also, over time, reduces the synaptic weight to avoid overloading the network (if all synapses were at maximum weight there would be many signals and little information passed); this is the process of forgetting.

Signals are also passed along neurons; for example, from the brain to the foot to control walking. These signals are transmitted as voltage spikes caused by a large flux of charged ions through protein pumps that traverse the neuronal cell membrane. Hodgkin-Huxley provided an electrical model for this system [24] where the cell membrane is modelled as a capacitor (the insulator layer is the lipid tails of phospholipids and the charged layer is the phosphate heads of the phospholipids (and decorated glycoproteins) and the protein pumps are modelled as time-varying non-linear resistors.

[3] The human brain contains a physiological reason why we can only perceive causal events and time in a forwards direction, despite results from physics suggesting that the laws of physics are time-direction-invariant.

1.3 Memristor Networks as Used in Artificial Intellegence

The memristor concept has proved useful in neuroscience. It has been shown that synapse action [27, 29] is memristive and the ion channels involved in neuronal signal transport are biological memristors, in fact, the time-varying resistors in Hodgkin-Huxley's electrical model of the neuron (as mentioned in subsection 1.2 have been more properly identified as memristors [8, 9].

In a computer neural net, the memristors can act as synapses where the conductivity of a memristor connection is taken as the synaptic weight of that connection. During the training period, the synaptic weight increases as the resistance of the memristor decreases. Some examples of memristive synapses in neural nets are [27, 33, 28, 11, 26]. As memristors are non-linear devices and the resistance-spike number curve is non-linear, the synapses are inherently non-linear, which introduces richer behaviour to the system. In [26] it was demonstrated that depending on the initial orientation of the curve, memristors can be included that are trained easily and quickly after few spikes, but slow their rate of change after many spikes (reducing the chance of network saturation) or memristors can be included to be slow to respond to low numbers of spikes and to change more drastically after larger numbers.

Memristor networks have only recently started to be investigated from the point of view of A.I. Other than neural nets, memristor networks have been shown (through simulation) to be useful to adding in memory to network dynamics [19, 18, 13].

An example [19] is a trained network that composed music where each possible pair of notes was connected by a pair of memristors (one each for each direction) which gave a fully connected directed graph. The network was trained by 'playing' music of a specific genre and altering connection weights between pairs of notes based on the amount that note combination was used. Tempo was included from a fully connected directed graph of note lengths. Once trained the network could compose music deterministically using Markov chain processes (Markov chains give the next note with only the knowledge of the previous note) where the memristors allowed network to maintain a memory beyond that allowed by standard Markov chain processes which is necessary for recognisable music complete with recurring themes and genre-specific style.

In another simple application [18], a matrix of light sensors attached to memristors acting under cellular automata rules was used to find the edges in an image. The memristor's short-term memory was used as a real-world cellular automata (where the automata's states were associated with the memristor losing it's short-term memory). This was then simulated for low quality, wide-angle distorted real world images of horizons to demonstrate that the system was capable of finding the horizon. This system is cheap, light-weight, low-power and composed of all-analogue electronics in hardware, which is suitable for robust tilt correction in Unmanned Aerial Vehicles, UAVs.

Models of memristor short-term memory have been used in simulation designs of scheduling algorithms inspired by ant behaviour [13]. In these algorithms, the

non-linear response of the memristor is used to model the decreasing freshness of data over time or the importance of sorted data over transmission. This was modelled for the collection of data from a large area network of sensors, but is applicable to sorted big data applications – important mission critical data is transferred first, allowing for quick decision making in an emergency, and nice-to-know useful data is transferred after allowing for slower, more considered data analysis when time allows. The algorithm was then applied to compressing images for quick transfer, where the most important part of the image (or data) could be sent first. Robot control systems that allowed the robot processor to switch on a continuum between fast, emergency decisions with only very relevant data (such as, is the robot's heat sensor too hot, indicating it is on fire) and considered learning and decisions (such as, how should the new information gained that day change the robots operation in the future).

1.4 Summary

In this chapter we present preliminary results from an ongoing investigation into if and how memristor spikes can affect neuronal cell spikes (i.e. can memristors act as an input to a living neuronal network) and if the spiking cells can alter the memristor dynamics – the latter question concerns us here. Various studies [15, 16, 20] have shown that the memristor spikes are reproducible and can interact through time on a single memristor and/or through space in a network of memristors. Here we show that neuronal cell spiking can affect the spiking properties of an attached memristor network.

The memristors used in this study [21] are flexible, operate with physiological currents [14], and are made of spun-on titanium dioxide sol-gel [17] between sputtered aluminium electrodes. We postulated that the current spikes could trigger spiking in cells, but it was not known if neuronal cell voltage spikes would affect the memristors: usually the lowest voltage the memristors are tested with is 0.01V and $I-V$ curves typically range over $\pm 1V$.

2 Methodology

The memristors used in this study have two different behaviours [21]. Previously we reported that using a mixture of both types connected either in series with opposite polarity or in an anti-parallel configuration was most likely to lead to spiking dynamics [16]. For this study, a circuit of two memristors of different types in anti-polarity series was used, see figure 2. The current was measured at 0.56s intervals for 560s using a Keithley sourcemeter (controlled via MatLab) that provides a voltage of +0.5V.

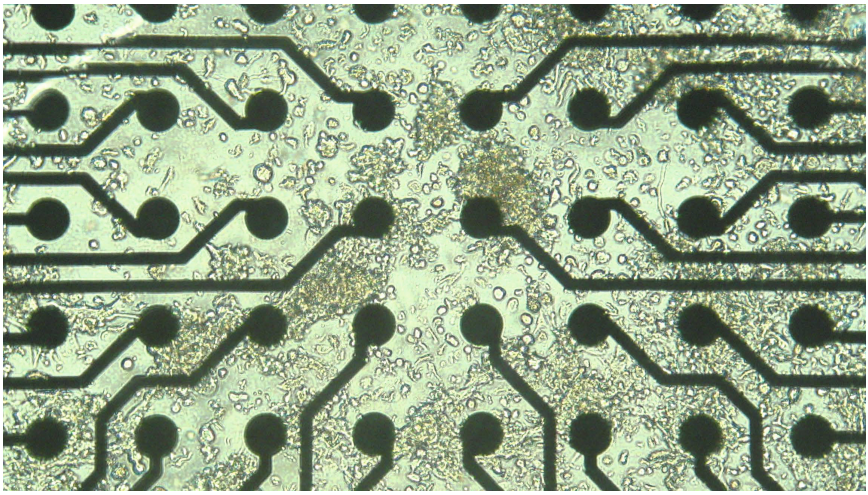
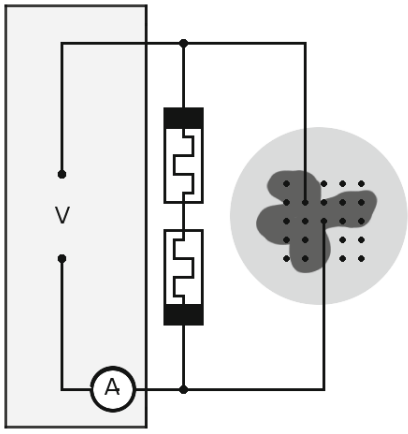


Fig. 1 IMR-32 cells in an MEA dish

Fig. 2 Scheme



The cells used in this study are IMR-32 neuroblastoma cells [37]. The cells were cultured in RPMI medium (RPMI-1640, Sigma-R0883) with 10% heat-inactivated fetal bovine serum (Sigma-F0804), 2mM L-glutamine (Sigma-G7513), essential amino acids (RPMI 1640 amino acids solution, Sigma-R7131) and 0.01% penicillin-streptomycin (Sigma-P1333). The cells (25,000 cells/dish) were transferred to a multi-electrode array (MEA) dish (Avenda Biosystems) coated with PEI [3] and incubated at physiological temperatures for between 24 and 48 hours. The MEA comprises 64 microelectrodes 30μM diameter and 200μM apart which allows the measurement of physiological voltages (typically $\leq$ 200mV). The two electrodes in close proximity of each other with the largest amount of spiking activity were selected to be connected to the external memristor circuit.

Two different ways of promoting cell spiking were used: 1. A subset of the IMR-32 cells were placed in a medium with added [dibutyryl cAMP, 0.5mg/ml, Sigma-D0627], which causes differentiation and promotes spiking [4]; 2. 10μL of 130mM capsaicin (Sigma-M2028) in phosphate buffered saline was added 200s into experiments with either undifferentiated, or differentiated cells - capsaicin is known to promote TRPV1-dependent calcium signalling, and differentiated IMR-32 cells have been shown to express TRPV1 [1]. The first dose of capsaicin was administered when the cells were not connected to the memristors to check that capsaicin caused increased spiking, then the memristors were connected for the second and third doses, further doses caused a negligible effect due to cell desensitisation.

Sixteen experiments will be discussed: 7 experiments with undifferentiated, unstimulated cells (with 3 control experiments comparing electrode separation) and 9 with stimulation of spiking by adding capsaicin, 3 of which were differentiated cells. For the memristor network control, the cables to the MEA dish were disconnected and left as an open-circuit. For the MEA dish controls, a clean MEA dish was filled with (in turn) distilled water, deionised water or RPMI medium as made up for the cells (and warmed to body temperature).

3 Results

3.1 Controls

3.1.1 Isolated Memristor Network

The response of the memristor network when not connected to the cells is shown in figure 3, this data was recorded after the cell spiking experiments however is qualitatively similar to that recorded beforehand (not shown) and in other spiking memristor networks [16]. This figure demonstrates expected 2-memristor network dynamics. A single memristor has a spike decay curve and a negative spike when the voltage is turned off, when combined into networks the starting (and ending) spikes are suppressed, instead the system shows sudden emergent bursting spikes, such as those seen after 200s in figure 3.

3.1.2 Memristor Network Connected to MEA Dish without Cells

Cell culture medium is an aqueous solution containing a mix of ions, sugars, colloids, fats and proteins, therefore any observed effect needs to be separated from that of connecting a memristor network to an aqueous sink of charged particles. Thus, we compared the results for an MEA dish filled with RPMI and distilled water. To completely remove the effects of ions on any conductivity, ultrapure deionised water was also tested.

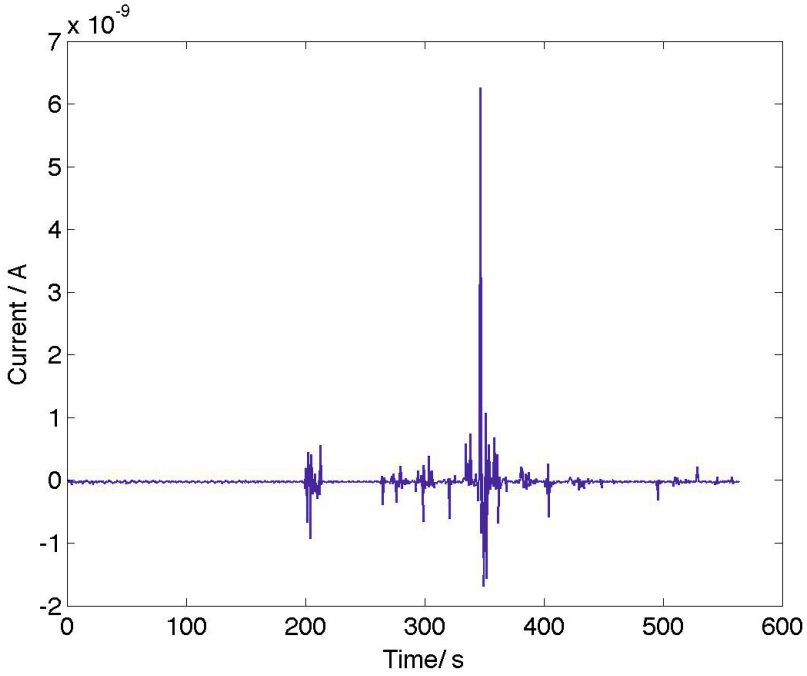


Fig. 3 Control Experiment 1: 2-memristor circuit without cells. Emergent spiking is seen after 200s

Results for the cell culture medium control are shown in 4 and this response is the expected response for a single memristor [15] or a memristor network with low compositional complexity and 'ideal' memristors. The other two controls (distilled water and deionised water) are identical and all start with the current of 60μA (a measured resistance of 8333Ω). The memristor spiking is suppressed and it appears that the current spikes are removed from the memristor network or otherwise damped by connection to a large sink of ions. The control experiments show that this effect is due to the loss of current from the networks to the aqueous medium, and not to cellular activity. Thus, it seems that connection to an external current sink removes the non-idealities of a memristor network (not permanently, because the network still exhibits standard memristor network dynamics after the experiments). We believe that the memristor network spiking is due to spike-response-caused voltage changes across individual memristors being out of sync (as postulated in [19]). If this is hypothesis is correct, then the connection to a sink of current would likely prevent this mechanism occurring by damping the system dynamics.

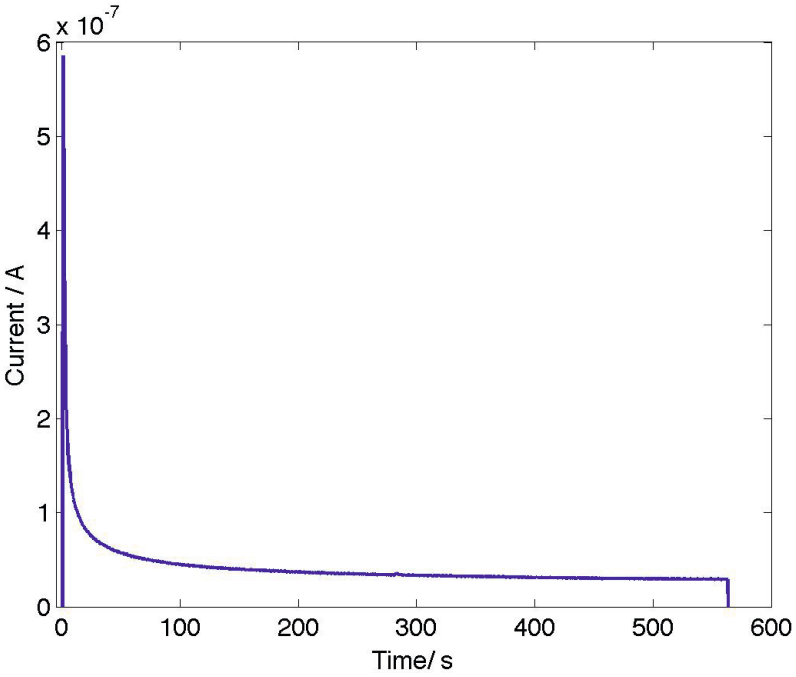


Fig. 4 Control Experiment 2: 2-memristor circuit connected to cell culture set-up containing cell medium but no cells

3.2 The Effect of Coupling to an External Pool of Spiking Neurons

Figure 5 shows the effect on the memristor $I-t$ profile of coupling the 2-memristor circuit to an MEA dish full of living, spiking neuronal cells. The memristor spiking is again suppressed, as with the control experiment involving medium. However, by comparing figure 5 with the control in figure 4, we can see that there is a difference due to the action of the cells: there is a discontinuity at 30s and a spike seen at 410s, as marked by arrows in figure 5. The discontinuity is similar to that seen when a voltage changes across the memristor and could be the result of a voltage spike changing the 'state' of the memristor, the spike may be caused by the cells themselves or a memristor spike retained in the network.

Figure 6 shows the same experiment for a different dish that contained more cells, had been incubated for an extra day and where the electrodes were closer (neighbours along the edges of the rectilinear grid rather than on the diagonal). This data clearly shows an effect of coupling to the cells at 300 and 400s. The spike at 400s is consistent to a memristor being switched to a higher voltage, and is suggestive of a cellular change altering the resistance of the memristor. The drop and spike at around 300s is not consistent with this, and could be the memristor network reacting to neuronal dynamics.

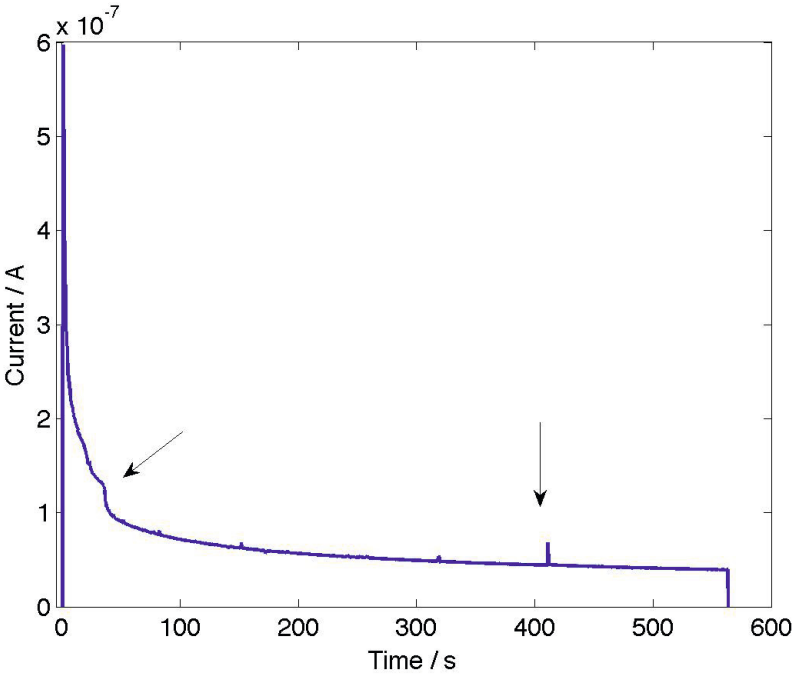


Fig. 5 The effect of cells. Measured $I-t$ curve for the 2-memristor circuit. Arrows indicate the possible effect of cellular processes

A full set of control experiments for electrode separation (not shown) have been run comprising of 1,2 and 3 diagonals, equivalent to 283μm, 566μm and 849μm, which demonstrate that the amount of perturbation of the memristor's $I-t$ curve is increased for shorter distances, i.e. the closer the electrodes, the shorter distance any signal has to travel to affect the exterior memristor circuit, thus the more likely it is that cellular processes will affect the memristors.

3.3 The Effect of Extra Spiking

Figure 7 show the effects of adding capsaicin to chemically stimulate cell spiking. The cells were subject to 10μL at approximately 10 minutes before starting the experiment and a second amount at 200s into this run. The main result is that increased cell activity has greatly changed the shape to a sloping background line with many spikes superimposed on it. This shows that an external memristor circuit can 'record' spiking activity from an electrically connected living neuronal network. The second result is that the effect of the increased capsaicin dose can be seen 110s after it was added in the form of an increased number and magnitude of the spikes.

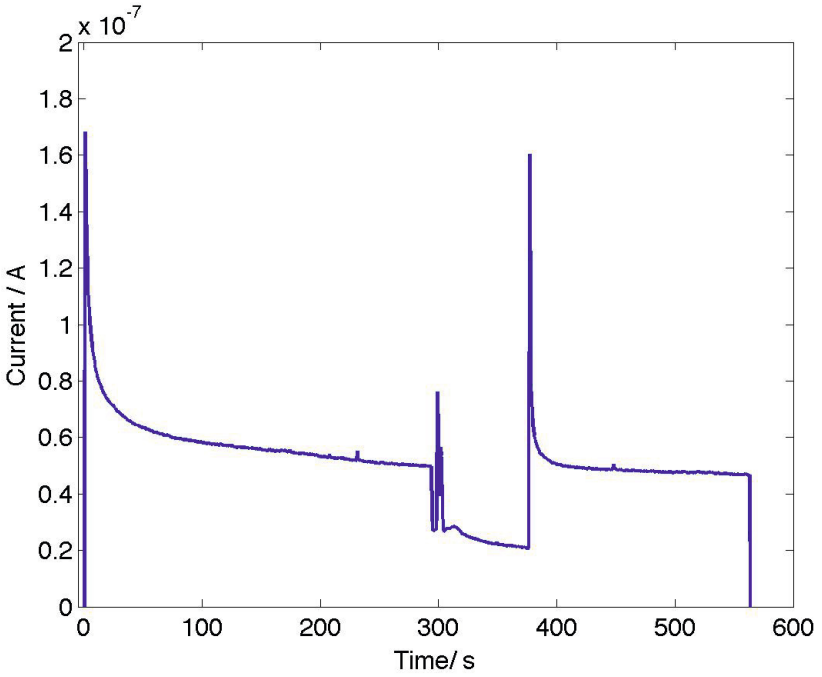


Fig. 6 Repeat of experiment, using a dish containing more cells and after an extra day's growth

3.4 Behaviours Observed

The types of behaviour were classified into 6 behaviours based on visual appearance:

'Curve no spikes' The expected single memristor response, as shown in figure 4
'Curve few spikes' The expected single memristor response with a small number of spikes superimposed

Table 1 The number of each type of graph from the preliminary results[a].

Behaviour	Curve no spikes	Curve few spikes	Curve with steps	Curve with switch	Line few spikes	Line many spikes
Total no.	2	2	4	1	1	6
Undifferentiated	2	0	1	1	1	2
With capsaicin	0	2	-	0	0	4
Differentiated*	0	0	3	0	0	0

[a] Differentiated results all had added capsaicin.

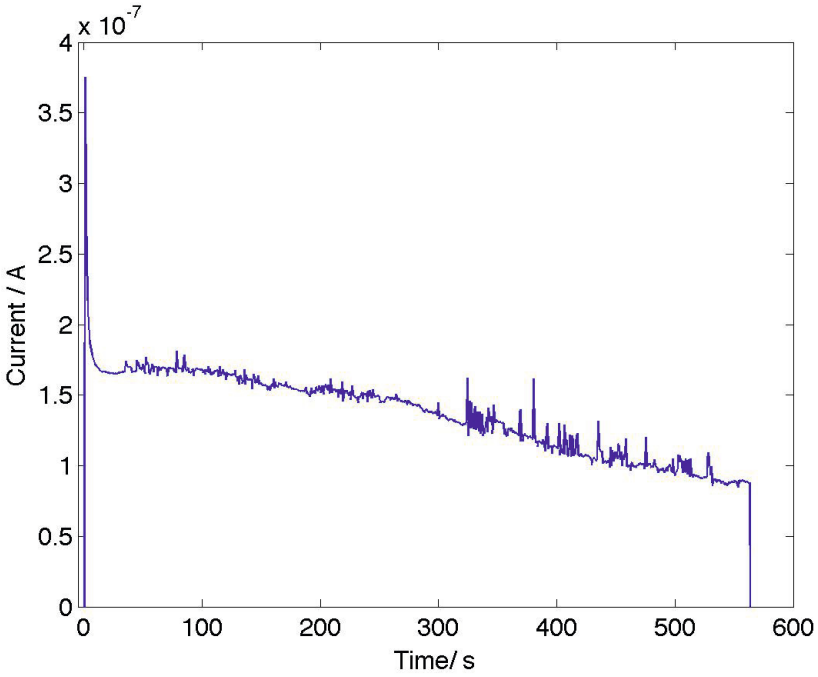


Fig. 7 Capsaicin added whilst memristors were connected

'Curve with steps' A curve with a change in gradient like the step shown in figure 5, multiple steps were seen in some cases
'Curve with switch' A change in memristor state, evidenced by an apparent change in the voltage applied to the memristor as in figure 6
'Line' Roughly sloping line
'Line many spikes' Roughly sloping line with many spikes as in figure 7

There are 16 individual data sets compared, which is enough to give an indication of an effect, but too few for statistical analysis. Bearing this in mind, we can see that the step superimposed with a curve is likely in differentiated cells (with a capsaicin-induced increase in spiking rate); conversely, the addition of capsaicin to undifferentiated cells is more likely to lead to a line with spikes. A standard curve with spikes overlaid was only seen under capsaicin, implying that these spikes come from the cells and not the memristors themselves (a capsaicin control was run and gave the same results for water). It is possible for some experiments to exhibit no spikes in the memristor circuit at all, even though the memristors were only connected to areas that were seen to be actively spiking, so in those cases cell spiking activity either stopped or was not transmitted to the memristor circuit. Finally, the cells with no chemical additives to induce spiking seem to have a range of behaviours as measured by the memristor sub-circuit.

4 Conclusions

This work demonstrates that memristors can be switched by neuronal cell action and we have shown that the memristor's state can be changed by cellular activity. This is an exciting result as this cellular input could be put into an artificial neural net (either one built from memristors or other components) allowing the A.I. to react to cellular processes.

Furthermore, we've shown that connecting spiking neuronal cells to a spiking memristor changes the memristor network dynamics in a qualifiable way, and related some of these changes to extra spiking activity. Thus memristors can be used as a biosensor for spiking activity and these sort of nanoscale networks placed in different parts of the brain could be inputs into wearable expert systems for disease monitoring or diagnosis.

Acknowledgements. E.G. would like to acknowledge the University of the West of England's SPUR award which funded this work.

References

1. Abdullah, H., Cosby, S.L., Heaney, L., McGarvey, L.: S123 the effect of rhinovirus infection on cough receptors on human sensory nerve and human primary bronchial epithelial cells. Thorax 66, A57 (2011)
2. Benninger, D.H., Lomarev, M., Lopez, G., Wassermann, E.M., Li, X., Considine, E., Hallett, M.: Transcranial direct current stimulation for the treatment of Parkinson's disease. Journal of Neurology, Neurosurgery & Psychiatry 81(10), 1105–1111 (2010)
3. Bull, L., Uroukov, I.S.: Initial results from the use of learning classifier systems to control in vitro neuronal networks. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation, pp. 369–376. ACM (2007)
4. Carbone, E., Sher, E., Clementi, F.: Ca currents in human neuroblastoma IMR32 cells: kinetics, permeability and pharmacology. Pflugers Archiv. 416(1-2), 170–179 (1990)
5. Caporale, N., Dan, Y.: Spike timing-dependent plasticity: a Hebbian learning rule. Annu. Rev. Neurosci. 31, 25–46 (2008)
6. Chabardés, S., Kahane, P., Minotti, L., Koudsie, A., Hirsch, E., Benabid, A.L.: Deep brain stimulation in epilepsy with particular reference to the subthalamic nucleus. Epileptic Disorders: International Epilepsy Journal with Videotape 4, S83–S93 (2002)
7. Chua, L.O.: Memristor - the missing circuit element. IEEE Trans. Circuit Theory 18, 507–519 (1971)
8. Chua, L., Sbitnev, V., Kim, H.: Hodgkin-huxley axon is made of memristors. International Journal of Bifurcation and Chaos 22, 1230011 (2012)
9. Chua, L., Sbitnev, V., Kim, H.: Neurons are poised near the edge of chaos. International Journal of Bifurcation and Chaos 11, 1250098 (2012)
10. Coiera, E.: The Guide to Health Informatics, 2nd edn. CRC Press (2003)
11. Ebong, I., Mazumder, P.: Memristor based STDP learning network for position detection. In: Proceesings of 2010 International Conference on Microelectronics (ICM), pp. 292–295 (2010), doi:10.1109/ICM.2010.5696142

12. Erokhin, V., Berzina, T., Camorani, P., Smerieri, A., Vavoulis, D., Feng, J., Fontana, M.P.: Material memristive device circuits with synaptic plasticity: learning and memory. BioNanoScience 1(1-2), 24–30 (2011)
13. Gale, E., de Lacy Costello, B., Adamatzky, A.: Memristor-based information gathering approaches, both ant-inspired and hypothetical. Nano Communication Networks 3(4), 203–216 (2012)
14. Gale, E.M., de Lacy Costello, B., Adamatzky, A.: Observation and characterization of memristor current spikes and their application to neuromorphic computation. In: 2012 International Conference on Numerical Analysis and Applied Mathematics (ICNAAM 2012), Kos, Greece (September 2012)
15. Gale, E., de Lacy Costello, B., Adamatzky, A.: Observation, characterisation and modeling of memristor current spikes. Appl. Math. Inf. Sci. 7, 1395–1403 (2013)
16. Gale, E., de Lacy Costello, B., Adamatzky, A.: Boolean logic gates from a single memristor via low-level sequential logic. In: Mauri, G., Dennunzio, A., Manzoni, L., Porreca, A.E. (eds.) UCNC 2013. LNCS, vol. 7956, pp. 79–89. Springer, Heidelberg (2013), doi:10.1007/978-3-642-39074-6_9
17. Gale, E., Mayne, R., Adamatzky, A., de Lacy Costello, B.: Drop-coated titanium dioxide memristors. Materials Chemistry and Physics 143, 524–529 (2014)
18. Georgilas, I., Gale, E., Adamatzky, A., Melhuish, C.: UAV Horizon Tracking Using Memristors and Cellular Automata Visual Processing. In: Natraj, A., Cameron, S., Melhuish, C., Witkowski, M. (eds.) TAROS 2013. LNCS, vol. 8069, pp. 64–75. Springer, Heidelberg (2014)
19. Gale, E., Matthews, O., de Lacy Costello, B., Adamatzky, A.: Beyond markov chains, towards adaptive memristors network-based music generation. International Journal of Unconventional Computing (2014)
20. Gale, E., de Lacy Costello, B., Adamatzky, A.: Emergent spiking in non-ideal memristor networks. Microelectronics Journal (2014), doi:10.1016/j.mejo.2014.06.008
21. Gale, E., Pearson, D., Kitson, S., Adamatzky, A., de Lacy Costello, B.: The effect of changing electrode material on solution-processed flexible titanium dioxide memrsitors. Materials Chemistry and Physics (2014) (Forthcoming)
22. Gergel-Hackett, N., Hamadani, B., Dunlap, B., Suehle, J., Richter, C., Hacker, C., Gundlach, D.: A flexible solution-processed memristor. IEEE Electron Device Letters 30(7), 706–708 (2009)
23. Hebb, D.O.: The organization of behavior: A neuropsychological theory. Psychology Press (2002)
24. Hodgkin, A.L., Huxley, A.F.: Propagation of electrical signals along giant nerve fibres. In: Proceedings of the Royal Society of London. Series B, Biological Sciences, pp. 177–183 (1952)
25. Howard, G., Gale, E., Bull, L., de Lacy Costello, B., Adamatzky, A.: Evolution of Plastic Learning in Spiking Networks via Memristive Connections. IEEE Transactions on Evolutionary Computation 16(5), 711–729 (2012)
26. Howard, G., Bull, L., de Lacy Costello, B., Gale, E., Adamatzky, A.: Evolution of Plastic Learning in Spiking Networks via Memristive Connections. Evolutionary computation 22(1), 79–103 (2014)
27. Jo, S.H., Chang, T., Ebong, I., Bhadviya, B.B., Mazumder, P., Lu, W.: Nanoscale memristor device as synapse in neuromorphic systems. Nano Letters 10(4), 1297–1301 (2010), doi:10.1021/nl904092h
28. Kim, H., Sah, M.P., Yang, C., Roska, T., Chua, L.O.: Memristor bridge synapses. Proceedings of the IEEE 100(6), 2061–2070 (2012), doi:10.1109/JPROC.2011.2166749

29. Linares-Barranco, B., Serrano-Gotarredona, T.: Memristance can explain spike-time-dependent-plasticity in neural synapses. Nature Precedings, 1–4 (2009)
30. Miranda, E.R., Bull, L., Gueguen, F., Uroukov, I.S.: Computer Music Meets Unconventional Computing: Towards Sound Synthesis with in vitro neuronal network. Computer Music 33, 9–18 (2009)
31. Patra, P.S.K., Sahu, D.P., Mandal, I.: An expert system for diagnosis of human diseases. Int. J. Comput. Appl. 13(1), 71–73 (2010)
32. Prodromakis, T., Chua, L.: Two centuries of memristors. Nature Materials 11, 478–481 (2012)
33. Querlioz, D., Bichler, O., Gamrat, C.: Simulation of a memristor-based spiking neural network immune to device variations. In: Proceedings of the 2011 International Joint Conference on Neural Networks (IJCNN). IEEE (2011)
34. Shortliffe, E.H.: Medical Expert Systems Knowledge Tools for Physicians. West. J. Med. 145(6), 830–839 (1986)
35. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. Nature 453, 80–83 (2008)
36. Szolovits, P., Patil, R.S., Schwartz, W.B.: Artificial Intelligence in Medical Diagnosis. Annals of Internal Medicine 108(1), 80–87 (1988)
37. Tumilowicz, J.J., Nichols, W.W., Cholon, J.J., Greene, A.E.: Definition of a continuous human cell line derived from neuroblastoma. Cancer Research 30(8), 2110–2118 (1970)
38. Uroukov, I.S., Ma, M., Bull, L., Purcell, W.M.: Electrophysical measurements in three-dimensional in vivo-mimetic organotypic cell cultures: Preliminary studies with hen embriyo brain spheroids. Neuroscience Letters 404, 33–38 (2006)

Author Index