



# Using emotions for the development of human-agent societies<sup>\*#</sup>

J. A. RINCON<sup>‡1</sup>, J. BAJO<sup>2</sup>, A. FERNANDEZ<sup>3</sup>, V. JULIAN<sup>1</sup>, C. CARRASCOSA<sup>1</sup>

(<sup>1</sup>Department of Computer Systems and Computation, Universitat Politecnica de Valencia, Valencia 46022, Spain)

(<sup>2</sup>Departamento de Inteligencia Artificial, Universidad Politecnica de Madrid, Madrid 28040, Spain)

(<sup>3</sup>Centre for Intelligent Information Technologies, Universidad Rey Juan Carlos, Madrid 28032, Spain)

E-mail: jrincon@dsic.upv.es; jbajo@fi.upm.es; alberto.fernandez@urjc.es; {vinglada, carrasco}@dsic.upv.es

Received Oct. 18, 2015; Revision accepted Feb. 17, 2016; Crosschecked Mar. 9, 2016

**Abstract:** Human-agent societies refer to applications where virtual agents and humans coexist and interact transparently into a fully integrated environment. One of the most important aspects in this kind of applications is including emotional states of the agents (humans or not) in the decision-making process. In this sense, this paper presents the applicability of the JaCalIVE (Jason Cartago implemented intelligent virtual environment) framework for developing this kind of society. Specifically, the paper presents an ambient intelligence application where humans are immersed into a system that extracts and analyzes the emotional state of a human group. A social emotional model is employed to try to maximize the welfare of those humans by playing the most appropriate music in every moment.

**Key words:** Multi-agent systems, Virtual environments, Emotional agents

<http://dx.doi.org/10.1631/FITEE.1500343>

**CLC number:** TP13

## 1 Introduction

Over the last few years different approaches have tried to develop ambient intelligence (AmI) applications based on intelligent systems. Nevertheless, the degree of success has not been as expected. There are two main reasons as the cause of this problem. First, intelligent systems have not reached the maturity level of other information technologies, and for a long time, they have forgotten traditional industry (Hendler, 2007). Second, an interdisciplinary per-

spective is required, which is hard to achieve, since a considerable number of available resources (scientific, economic, and human) would be required.

Agent technology, although still immature in some ways, allows the development of systems that support the typical requirements of AmI applications. Specifically, agent technology allows the creation and management of systems where the main components can be humans and software agents providing services to humans or other agents in an environment of whole integration. This kind of applications is what we call a human-agent society (Billhardt *et al.*, 2014), which can be defined as a computing paradigm in which the traditional notion of application disappears. Instead of developing software applications that accomplish computational tasks for specific purposes, this paradigm is based on an immersion of the users in a complex environment that

<sup>‡</sup> Corresponding author

\* Project supported by the Ministerio de Economía y Competitividad of the Spanish Government and the European Regional Development Fund of the European Union (No. TIN2015-65515-C4-1-R)

# A preliminary version was presented at the 13th International Conference on Practical Applications of Agents and Multi-Agent Systems, June 3–4, 2015, Spain

ORCID: V. JULIAN, <http://orcid.org/0000-0002-2743-6037>

© Zhejiang University and Springer-Verlag Berlin Heidelberg 2016

enables computation. Nevertheless, working with humans is complex, because they use emotions in their decision making. Human beings manage themselves in different environments, either in the working place, at home, or in public places. At each one of these places we perceive a wide range of stimuli, which interfere in our commodity levels modifying our emotional levels. These variations in our emotional states could be used as very useful information for software agents. Nevertheless, it is required that the agents have the capability of interpreting or recognizing such variations. This is the reason for implementing emotional models that interpret or represent the different emotions.

Our proposal is to employ the emotional state of a group of agents (humans or not) in an AmI application. Concretely, we propose in this paper a system for controlling automatically the music which is playing in a bar. The main goal of the disc jockey (DJ) is to play music making that all individuals within the bar are mostly as happy as possible. Each of the individuals will be represented by an agent, which has an emotional response according to his/her musical taste. That is, depending basically on the musical genre of the song agents will respond varying their emotional states. Moreover, varying emotions of each agent will modify the social emotion of the group. The application has been developed using the JaCalIVE (Jason Cartago implemented intelligent virtual environment) framework (Rincon *et al.*, 2014), which is a framework for the design and simulation of intelligent virtual environments (IVEs). This framework differs from other works in the sense that it integrates the concepts of agents, humans, artifacts, and physical simulation. Besides, IVEs developed using the JaCalIVE framework can be easily modified thanks to Xtensible Markup Language (XML) modelling and automatic code generation. The main reason to employ JaCalIVE is that it allows an easy integration of human beings in the system. This framework can be downloaded from <http://jacalive.gti-ia.dsic.upv.es>.

## 2 Related work

Human-agent societies are an evolution and integration of different paradigms and technologies that have appeared in the last decades. In the first

place, they are an evolution of multi-agent systems (MAS) which have been open systems (Zambonelli *et al.*, 2003) characterized by their limited confidence, possible individual goals in conflict, and a great probability of non-accordance with specifications. Current MAS approaches try to solve these conflicts using aspects like norms, organizations, negotiation, argumentation, coordination, and trust (Ossowski, 2013). These new technologies can be considered the building blocks for human-agent societies, but they have to be enhanced beyond the computational space for the development of socially acceptable intelligent systems in which humans represent just another component. In the second place, human-agent societies are based on the developments in the field of service-oriented computing (SOC) and semantic technologies. SOC is a field that has evolved in the past in parallel to MAS, but both have a principal aspect in common: the idea of computing by delegation. Some approaches in this sense have already been proposed in the context of service-oriented multi-agent systems (Huhns *et al.*, 2005; Fernandez and Ossowski, 2011). Finally, the third pillar for constructing human-agent societies arises from the field of ambient intelligence.

AmI (Satyanarayanan, 2002; Mangina *et al.*, 2009) changes the concept of smart home, introducing new devices that help improve people's life quality, devices that learn our tastes, smart homes that help reduce energy consumption (Han and Lim, 2010), safer homes for elderly (Satyanarayanan, 2001; Intille, 2002), among other applications. To achieve this, AmI and ubiquitous computing employ different artificial intelligence tools, sensor networks, mobile Internet connections, and new and sophisticated embedded devices. AmI imagines a future where technology surrounds users, and helps them in their daily lives. The AmI scenarios described by the Information Society Technologies Advisory Group (ISTAG) exhibit intelligent environments capable of recognizing and responding to the presence of different individuals in a simple, non-intrusive, and often unseen way (Ducatel *et al.*, 2001). AmI is heavily based on the concept of ubiquitous computing (Weiser, 1991), which describes a world where a multitude of computational objects communicate and interact in order to help humans in daily activities. The main aim of AmI systems is to be invisible,

but very useful. This raises three requirements for AmI based systems (Satyanarayanan, 2001) (all of which are also relevant for human-agent societies): (1) the technology should be transparent to users; (2) the services must be adapted to the context and user preferences; (3) applications must provide intuitive and user-friendly interfaces. These kinds of systems represent an immense and ever-growing multidisciplinary area of research and development. It brings a huge technological innovation and impact for the citizens and society as a whole. Because the AmI technology is integrative, it has recently grown along with various disciplines including sensing technologies, wireless networking, software products and platforms, artificial intelligence, data analytics, and human-computer interfaces. We can detect a lack of research on how to use the existing technology to the best possible effect. The automatic recognition of human activities (i.e., emotions) and abnormal behaviours is an obvious prerequisite for new AmI applications, and requires novel methods to improve recognition rates, enhance user acceptance, and preserve the privacy of monitored individuals. These are challenging issues which must be addressed.

Similar to AmI and ubiquitous computing, the main challenge to achieve real human-agent societies lies in the design and construction of intelligent environments, in which humans interact with autonomous, intelligent entities, through different input and output devices. This means that there are two layers in which humans interact within the environmental and ubiquitous computing intelligence. The first layer is the real world where humans interact with other humans and with real objects. The second layer is a virtual layer in which humans interact with virtual entities and objects. The latter layer will be inhabited by intelligent entities (agents), which must be able to perform the different human orders. These virtual environments where agents are involved, are known as IVEs.

An IVE (Hale and Stanney, 2002) is a 3D space that provides the user with collaboration, simulation, and interaction with software entities, so he/she can experience a high immersion level. This immersion is achieved through detailed graphics, realistic physics, artificial intelligence (AI) techniques, and a set of devices that obtain information from the real world. The JaCaIIVE framework enables the

design, programming, and deployment of systems of this kind. In these IVEs different entities exist, which perform specific tasks. The framework facilitates to find agents in charge of accessing databases, agents which control some kind of complex objects, and agents which represent humans. These agents are in charge of serving as a wrapper between the real world and the virtual world. They help humans interact with other virtual entities, which can be representations of other humans or entities performing some control in the real world. This allows humans a transparent interaction in both real and virtual environments. To allow agents to interact with the real world, it is necessary that these agents have access to specific devices which allow to collect real-world information. Devices such as cameras, Kinect, and microphones allow agents to perceive the environment improving their interaction. In this sense, Section 4 introduces how to use the proposed JaCaIIVE framework.

Regarding the simulation of emotions in multi-agent systems, we can find two relevant emotional models: the OCC model (Ortony, 1990) and the PAD model (Mehrabian, 1997). These emotional models are the most used ones to detect or simulate emotional states. The OCC model classifies human emotions into 22 categories which are divided into five processes: (1) the classification of events, actions, or objects found; (2) the intensity quantification of the affected emotions; (3) the association of interactions between the generated emotion with the existing ones; (4) cartography of the emotional state according to the emotional expression; and (5) the expression of the emotional state through facial expressions (Ortony, 1990). These processes define the whole system, where the emotional states represent the way of perceiving our environment (objects, persons, places) and, at the same time, influencing positively or negatively in our behaviours (Ali and Amin, 2014). The OCC model provides a good starting point to integrate an emotional model into an intelligent software entity. However, the OCC model utilization and implementation present one important problem due mainly to its high dimensionality.

In this sense, the PAD model is a simplification of the OCC model. It allows the representation of the emotion in  $\mathbb{R}^3$ . Each one of the components conforming this emotional model, allows to represent a

measure of an emotional state. In this way, we obtain a numerical representation of all the emotions (Bales, 2001). The three employed values are usually normalized in the range  $[-1, 1]$ , and correspond to the three components conforming the emotional model (pleasure, arousal, dominance).

The pleasure-displeasure ( $P$ ) scale measures how pleasant an emotion may be. For instance, both anger and fear are unpleasant emotions, and score high on the displeasure scale. However, joy is a pleasant emotion. This dimension is usually limited to 16 specific values (Mehrabian, 1980). The arousal-nonarousal ( $A$ ) scale measures the intensity of the emotion. For instance, while both anger and rage are unpleasant emotions, rage has a higher intensity or a higher arousal state. However, boredom, which is also an unpleasant state, has a low arousal value. This scale is usually restricted to 9 specific values (Mehrabian, 1980). The dominance-submissiveness ( $D$ ) scale represents the controlling and dominant nature of the emotion. For instance, while both fear and anger are unpleasant emotions, anger is a dominant emotion, and fear is a submissive emotion. This scale is, the same as the previous one, usually restricted to 9 specific values (Mehrabian, 1980).

The existing emotional models are thought to detect and/or simulate human emotions but only for a lonely entity. That is, it is not taken into account the possibility of representing multiple emotions inside a heterogeneous group of entities, where each one of such entities has the capability of detecting and/or emulating one emotion. Due to this, the next section presents an approximation of a social emotional model based on PAD. This new model allows to detect the emotion of a group of entities.

### 3 Social emotion

The social emotion (SE) model (Rincon *et al.*, 2015a; 2015b) allows obtaining the emotion of a group of heterogenous agents. The model also allows to obtain the emotional distance between the emotion of a group and a target emotion. This model is based on the PAD model, which is employed as a way to represent the emotion of an agent  $ag_i$ . This emotion is represented as a vector in  $\mathbb{R}^3$ , represented by three components that make up the PAD emotional

model:  $P$ =pleasure,  $A$ =arousal, and  $D$ =dominance. The variation of each component allows to modify the emotional state of the agent:

$$\mathbf{E}(ag_i) = [P_i, A_i, D_i]. \quad (1)$$

Once defined how the emotional representation of each agent is, the next step consists of calculating the central emotion (CE) for a group of agents  $Ag = \{ag_1, ag_2, \dots, ag_n\}$ . This representation is obtained by averaging their  $P, A, D$  values:

$$\bar{P} = \frac{\sum_{i=1}^n P_i}{n}, \bar{A} = \frac{\sum_{i=1}^n A_i}{n}, \bar{D} = \frac{\sum_{i=1}^n D_i}{n}. \quad (2)$$

The result of Eq. (2) is also a vector in  $\mathbb{R}^3$  which is the core emotion ( $\mathbf{CE}(Ag)$ ) of a group of agents:

$$\mathbf{CE}(Ag) = [\bar{P}, \bar{A}, \bar{D}]. \quad (3)$$

However, this element by itself is, obviously, not enough to obtain a correct measure of the social emotion of a group of agents. It is necessary to introduce more information about the possible cohesion of each agent inside the group. This new component is the maximum distance between  $\mathbf{CE}(Ag)$  and the PAD value of each agent ( $\mathbf{m}(Ag)$ ), where for  $\mathbf{m}(Ag)$  each component is calculated by means of the Euclidean distance of each value of the PAD (Eqs. (4)–(6)) as follows:

$$m_P(Ag) = \max(|P_i - \bar{P}(Ag)|), \forall ag_i \in Ag, \quad (4)$$

$$m_A(Ag) = \max(|A_i - \bar{A}(Ag)|), \forall ag_i \in Ag, \quad (5)$$

$$m_D(Ag) = \max(|D_i - \bar{D}(Ag)|), \forall ag_i \in Ag, \quad (6)$$

$$\mathbf{m}(Ag) = [m_P(Ag), m_A(Ag), m_D(Ag)]. \quad (7)$$

The  $\mathbf{m}(Ag)$  lets us know which agent is the farthest from the central emotion of the group. However, in order to have an idea about the possible dispersion of the agents around the central emotion, the standard deviation (SD) is introduced in the model. So, the SD allows to calculate the level of emotional dispersion of a group of agents around the central emotion ( $\mathbf{CE}(Ag)$ ) and it is calculated for each

component of the PAD:

$$\left\{ \begin{array}{l} \sigma_P(\text{Ag}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - \bar{P}(\text{Ag}))^2}, \forall \text{ag}_i \in \text{Ag}, \\ \sigma_A(\text{Ag}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (A_i - \bar{A}(\text{Ag}))^2}, \forall \text{ag}_i \in \text{Ag}, \\ \sigma_D(\text{Ag}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (D_i - \bar{D}(\text{Ag}))^2}, \forall \text{ag}_i \in \text{Ag}. \end{array} \right. \quad (8)$$

The results of these equations are collected as a vector (Eq. (9)), which allows to determine the degree of emotional dispersion of the group:

$$\boldsymbol{\sigma}(\text{Ag}) = [\sigma_P(\text{Ag}), \sigma_A(\text{Ag}), \sigma_D(\text{Ag})]. \quad (9)$$

In this case, if  $\boldsymbol{\sigma}(\text{Ag}) \gg [0, 0, 0]$ , the group has a high emotional dispersion; i.e., the members of the group have different emotional states. On the other side, if  $\boldsymbol{\sigma}(\text{Ag}) \cong [0, 0, 0]$ , the group has a low emotional dispersion, which means that the agents have similar emotional states.

Taking into account the three presented elements, the social emotion (SE) of a group of agents  $\text{Ag} = \{\text{ag}_1, \text{ag}_2, \dots, \text{ag}_n\}$  can be defined by the following triplet:

$$\text{SE}(\text{Ag}) = (\mathbf{CE}(\text{Ag}), \mathbf{m}(\text{Ag}), \boldsymbol{\sigma}(\text{Ag})). \quad (10)$$

This representation also allows the calculation of emotional distances among different groups of agents or between a group of agents and a target emotion. This calculation can be used as a feedback in the decision-making process in order to do actions which try to move the social emotion to a particular area of the PAD space. For instance, agents can try to move the social emotion of an agent group to the happiness state.

Once the social emotional model has been presented, the next section introduces the development framework employed in the design and implementation of the proposed application.

## 4 JaCalIVE

In the last years, there have been different approaches for using MAS as a paradigm for modelling

and engineering IVEs, but they have some open issues: low generality and then reusability, weak support for handling full open and dynamic environments where objects are dynamically created and destroyed.

As a way to tackle these open issues, and based on the MAM5 meta-model (Barella *et al.*, 2012), the JaCalIVE framework was developed (Rincon *et al.*, 2014). This framework provides a method to develop this kind of applications along with a supporting platform to execute them. The presented work has extended the MAM5 meta-model along with the JaCalIVE framework to develop human-agent societies, that is, to include humans in the loop.

MAM5 allows to design an IVE as a set of entities that can be virtually situated or not. These entities are grouped inside workspaces (IVE workspaces in the case of virtually situated). Entities are divided into artifacts (IVE artifacts if situated) and agents (inhabitant agents if situated). One new type of situated agents, i.e., inhabitant agents are human-immersed agents, which model the human inside the system. Fig. 1 presents the extended MAM5 meta-model.

Fig. 2 shows the steps that should be followed in order to develop an IVE according to the JaCalIVE framework:

1. Model: The first step is to design the IVE. JaCalIVE provides an XML Schema Definition (XSD) based on the MAM5 meta-model. According to it, an IVE can be composed of two different types of workspaces depending on whether they specify the location of its entities (IVE\_Workspaces) or not (Workspaces). It also includes the specification of agents, artifacts, and the norms that regulate the physical laws of the IVE workspaces.

2. Translate: The second step is to automatically generate code templates from design. One file template is generated for each agent and artifact. JaCalIVE agents are rational agents based on Jason. The artifacts representing the virtual environment are based on CArtAgO. The developer must complete these templates and then the IVE is ready to be executed.

3. Simulate: Finally, the IVE is simulated. As shown in Fig. 2, the JaCalIVE platform uses Jason (<http://jason.sourceforge.net/wp/>), CArtAgO (<http://cartago.sourceforge.net/>), and

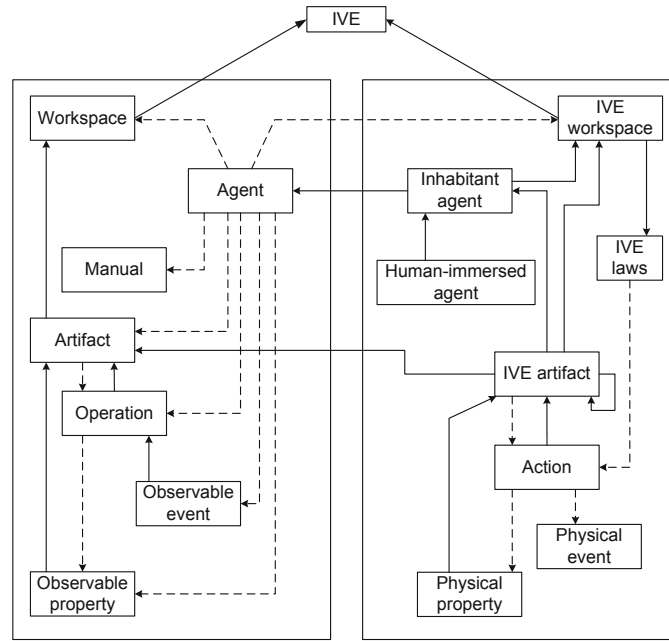


Fig. 1 Extended MAM5 to include human-immersed agents

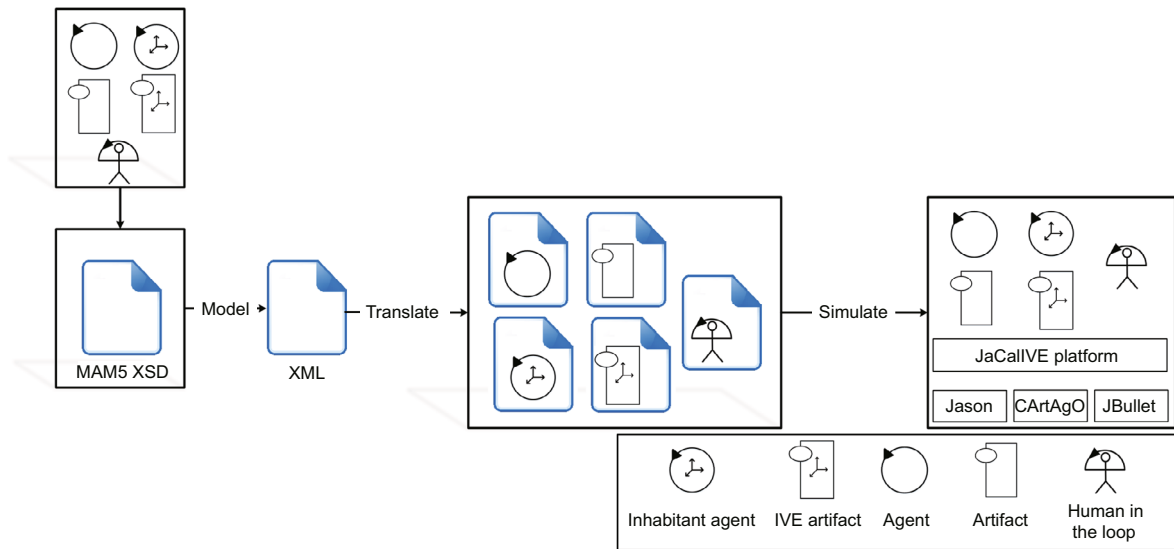


Fig. 2 General scheme, JaCaIVE

JBullet (<http://jbullet.advel.cz>). Jason offers support for BDI agents that can reason about their beliefs, desires, and intentions. CArtAgO offers support for the creation and management of artifacts. JBullet offers support for physical simulation. The JaCaIVE platform also includes internal agents (Jason based) to manage the virtual environment.

In summary, this framework differs from other works in the sense that it integrates the concepts of agents, artifacts, humans, and physical simulation. Besides, IVEs developed using the JaCaIVE framework can be easily modified thanks to the XML modelling and the automatic code generation. Following the MAM5 perspective, the modules used to

interact with the developed IVEs are uncoupled from the rest of the system. It allows to easily integrate different kinds of modules as needed. For example, it allows to adapt the visualization render to the requirements of the specific IVE we want to simulate. A more detailed description of the JaCalIVE framework can be found in Rincon *et al.* (2014).

## 5 Case study

### 5.1 Problem description

The application presented in this paper is centered in the analysis of the emotional state of a group of people trying to improve their emotional state through the use of music and taking into account the number of people around them as a way to influence the human mood. Concretely, the example has been developed in a bar, where there is a DJ in charge of playing music and a specific number of persons listening that music. At the same time, a person perceives how many people are around him/her and listen, with too much or too little interest, the music played by the DJ. According to his/her perceptions each one of the individuals placed in the bar, can modify his/her emotional state. The main goal of the DJ is to play music making that all the people within the bar are mostly as happy as possible. Moreover, people can move around the bar looking for positions where there are more or fewer people. This behaviour can also affect the emotional state of the individuals, and also, of the group.

In a specific moment, each one of the persons placed in the bar will have an emotional response according to his/her musical taste and if it has more or fewer agents around him/her. That is, the emotional state of the agent depends on the song that is playing and the number of agents that are within a comfort radius. People will respond varying their emotional states. Then varying emotions of each person will modify the social emotion of the group.

In such a way, the proposed application seeks to identify the different emotional states using them as a tool of communication between humans and agents. To perform this detection, we need to use pattern recognition algorithms and image and audio processing techniques in order to detect and classify the different emotional states of humans and try to mod-

ify the environment in which the humans are. The application has been developed as a virtual multi-agent system using the JaCalIVE framework where there will be different entities which will have specific roles. Each of these entities may represent real human beings or software agents that simulate humans, for instance, the DJ. The main characteristics of the proposed agents are defined in the following subsection.

### 5.2 Application design

The application is composed of three types of entities: (1) SEtA, which is responsible for calculating the social emotions of the agent group, (2) a DJ agent which is responsible for playing the music, and (3) different agents that represent humans who are inside the bar (Fig. 3).

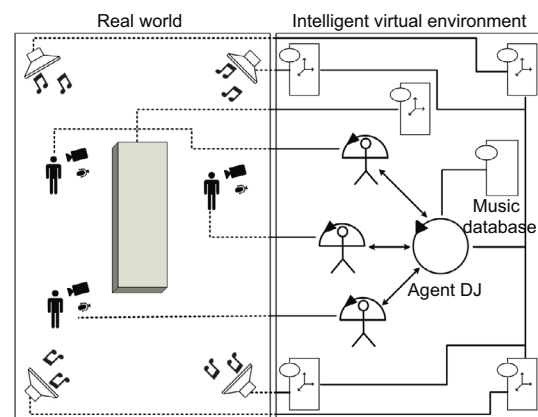


Fig. 3 Graphical view of the proposed application

The main tasks of these agents are:

1. SEtA agent: It is in charge of calculating the social emotion of a group of agents. This agent contains the model that was explained in Section 3. The aim of this agent is to obtain the emotion of all agents that live in the environment. Using this information the agent is capable of calculating the social emotion.

2. DJ agent: It is in charge of selecting and playing music in the bar. The main goal of this agent is to achieve an emotional state of happiness for all of the people in the bar. When the DJ agent plays a song, it must analyze the emotional state of people. According to this analysis, it will select the most appropriate songs in order to improve, if possible, the current emotional state of the audience.

To do this, the DJ agent evaluates the information given by the agent SEtA. Using this information it can know the effect that the songs have over the audience. This will help the DJ agent decide whether to continue with the same musical genre or not, in order to improve the emotional state of the group.

3. Human-immersed agent: It is in charge of detecting and calculating the emotional state of an individual in the bar. This information will be sent to the DJ agent using a subscription protocol. In order to accomplish its tasks, this agent must have access to a variety of input/output information devices as cameras, microphones, etc.

In order to facilitate the access to this kind of device, the devices have been modelled as artifacts (Fig. 3—in this figure and in the following ones, we used data from a simple example with only three human-immersed agents). Concretely, an artifact has been designed for managing each camera which allows face detection; each microphone is managed by an artifact which captures the ambient sound in order to classify the music genre; the music database has been designed as an artifact employed by the DJ agent (it stores around 1000 songs classified by genres) and there is an artifact for controlling the multimedia player and the amplifiers for playing songs in the bar with the appropriate volume.

Each one of the commented entities has been designed using JaCalIVE through an XML file describing all its different properties (including physical ones). The XML file allows to describe if you need some kind of sensor to capture information, some type of actuator or simply an agent that does not need real-world information. It is also in this XML file where humans are associated with each

agent. These XML files are automatically translated into code templates using the JaCalIVE framework.

### 5.3 Implementation

The proposed application has three different types of agents, as above commented: the SEtA, the DJ agent, and the human-immersed agent. Due to space limitations, this section focuses mainly on the human-immersed agent.

The human-immersed agent, is an agent that will monitor the associated person and will count the number of persons within his/her comfortable radius. This number and the musical profile allow the calculation of the emotional state of each human presented in the bar. Fig. 4 gives an overview of the different processes done by the human-immersed agent.

The first process is responsible for capturing information from the real world. This information is obtained by the human-immersed agent using an Asus Xtion and a microphone.

The second process is responsible for extracting the most relevant information for emotion detection, using the different images and ambient sounds. This information allows the human-immersed agent to make two things. The first one is a face detection process using the Viola-Jones algorithm (Viola and Jones, 2004). As a second thing, the human-immersed agent uses these images to identify the nearby humans.

The images detected and extracted using the Viola-Jones algorithm are resized to  $92 \times 112$  pixels. Each one of these matrices is transformed to a uni-dimensional vector. This vector is formed by concatenating the rows of the image matrix. So, for

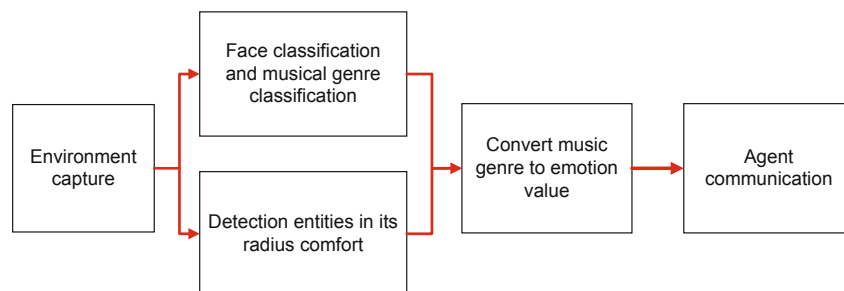


Fig. 4 Steps followed by the human-immersed agent

a  $92 \times 112$  image it will have 10 304 elements. All the image vectors are grouped forming a matrix (each vector is one row) of  $N \times M$ , where  $N$  represents the number of images and  $M$  is the maximum number of pixels of the initial images. Clearly, this new matrix is very big and the number of values is more than what is required (Fig. 5).

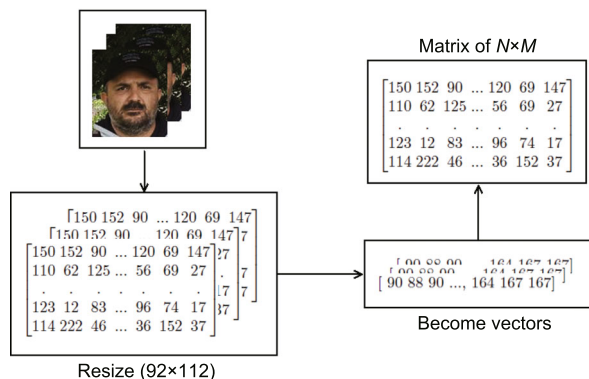


Fig. 5 Image recognition process

Once you have the  $N \times M$  matrix with all the images of our data set, it is necessary to reduce the dimensionality. The principal component analysis (PCA) algorithm allows us to perform this task, reducing our dimensionality as can be checked in Table 1.

Table 1 Comparison between applying or not PCA to image processing

	Number of images	Number of pixels
Without PCA	30	10 304
With PCA	30	2

PCA: principal component analysis

This new matrix is used with the K-means algorithm to partition it into  $k$  groups (Fig. 6). This is a fast method to determine if the face seen by the agent is the same as what has been associated to him/her. Nevertheless, this does not mean you cannot use other more complex face recognition algorithms such as support vector machines (Osuna et al., 1997) or neural networks (Lawrence et al., 1997), among others.

Moreover, the human-immersed agent captures the ambient sound in order to classify the music genre. To make it possible the human-immersed agent uses its microphone to capture the ambient

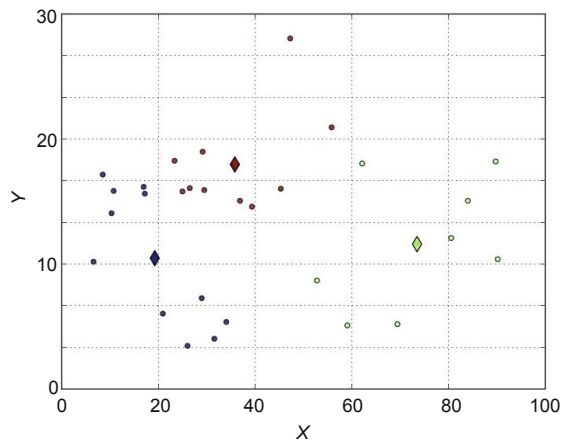


Fig. 6 Face clusters identified in the example of the application. Each dot represents a face and a diamond represents a group of similar faces. Different colors represent different groups. References to color refer to the online version of this figure

sound. There are different possibilities in the literature to classify music (Talupur et al., 2001; Li et al., 2003). The human-immersed agent uses a statistical classifier (Holzapfel and Stylianou, 2007) to decide which musical genre fits with the input song from 10 different genres: blues, country, electronic music, funky, heavy metal, pop music, rock, soul, pop-rock, and tropical music.

The third process uses the information obtained by the musical genre classifier to obtain the new emotional value. This emotional value is obtained using a fuzzy logic algorithm, which returns three values corresponding to the PAD model (Mehrabian, 1997; Nanty and Gelin, 2013). To obtain these values, it is necessary to know how the different musical genres influence the human. This information allows to modify the membership function of the fuzzy logic algorithm. The variation of this membership function depends on the corresponding human musical preferences; e.g., a person can respond favorably to pop music but not to blues music. This is the reason why it is required that each person should configure its human-immersed agent before using the system, varying the membership function of the fuzzy logic module.

However, this information is not enough to calculate the current emotional state of an individual. This is because humans have different reactions according to the number of persons placed around

them. So, each human-immersed agent has defined a circle of comfort. In this circle, the agent allows a limited number of other agents. This number can be different for each agent. According to this, if the number of nearby agents is greater than a threshold, the agent will modify its emotional state in a negative way.

Regarding the DJ agent, it will play music obtained from a playlist created in Spotify (<https://www.spotify.com>) using Mopidy (<https://www.mopidy.com/>) as communication API. This agent uses the information sent by the human-immersed agent to analyze the group's emotional state, using it to decide which is the most appropriate next song to play. The goal of the DJ agent is that all the humans feel as happy as possible.

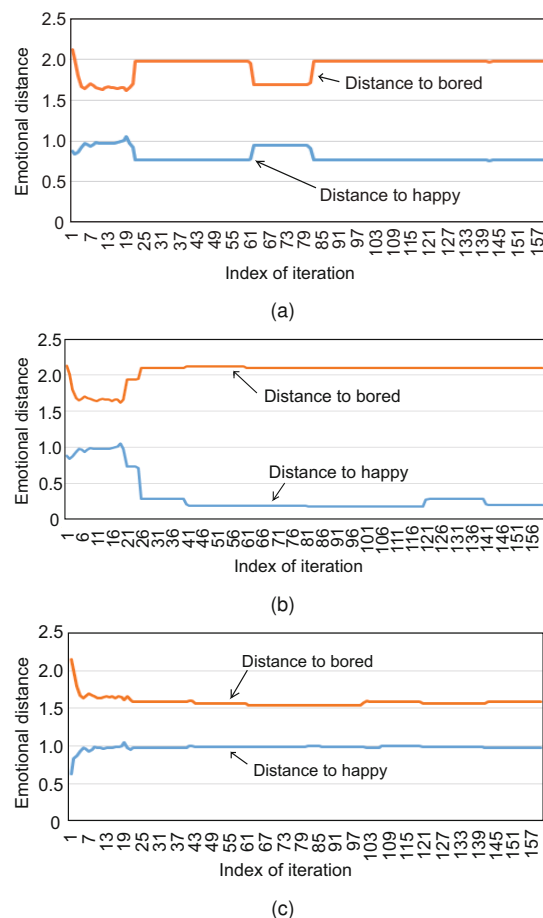
Different experiments have been developed in order to validate the proposed example. Specifically, the aim of these tests is to validate the use of the social emotion of a group of agents as a way to improve the decision-making process. Moreover, the experiments allow us to confirm the utility of the integration of agents, artifacts, humans, and physical simulation using the JaCalIVE framework. The experiments show the effect produced in the simulation when the music played in the bar is changed over time and also the number of people in the bar is increased.

## 6 Experiments

To do this, the implemented prototype has been tested changing the music played in the bar (around 156 songs are played in an iterative way). The experiments have been evaluated with different numbers of individuals within the bar: Low,  $\leq 20$ ; Medium, 20–80; High,  $> 80$ . All the tests were composed by agents representing people in the bar. These agents have assigned random initial emotions at the beginning of the execution of each test. Each agent responded differently to each song by means of a fuzzy logic algorithm that allows to change the emotional answer. Each emotional state of each agent was then used to calculate and to store the social emotion of the group. The dispersion and maximum distances of such emotional values with respect to the target emotions were also stored. As commented above,

the aim of the system was to minimize the distance between current social emotion and the target emotion, that is, happiness. So, the DJ agent would try to minimize such distance playing different songs of different genres according to the evolution of the social emotion.

Fig. 7 shows the results obtained regarding the distance between the social emotion of the group and the emotions happy and bored. As we can see, changing the number of people causes an emotional reaction which considerably affects the distance to happiness or to boredom. In this case, the minimum distance to happiness is obtained when the number of people is neither high nor low. In the experiment where the number of people is high, we can see how the social emotion maintains a distance against the two possible target emotions. In this sense, all the



**Fig. 7 Experiment 1: evaluation of the distance between the group emotion against the emotions of happiness and boredom. The number of people is low ( $\leq 20$ ) (a), medium (20–80) (b), or high ( $> 80$ ) (c)**

agents can have similar or very dissimilar emotions. This situation can be analyzed in more detail analyzing the maximum distances and the dispersion of the social emotion. These aspects were tested in the following evaluations.

In the same way, Fig. 8 presents the evolution of the maximum distances of any member of the group against happiness. As commented above, this experiment allows to observe if there exist agents having their emotional state far away from the central emotion. This information can be used by the system to try to act over these agents in order to reduce their emotional distance or even to exclude these agents in order to have a more cohesive group. In the figure, it can be observed that the distances become very stable as time passes. That is, there are some people

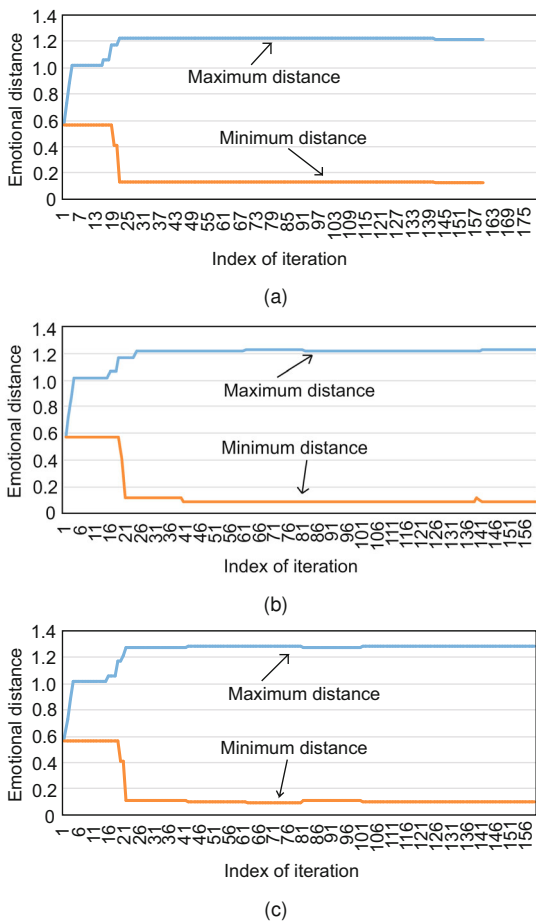


Fig. 8 Experiment 2: evaluation of the maximum and minimum distances between the emotion of some member of the group against happiness. The number of people is low ( $\leq 20$ ) (a), medium (20–80) (b), or high ( $> 80$ ) (c)

who maintain their emotional state away from happiness independent of the number of people in the bar. This situation can be due by a few agents or maybe by a large number of agents. This can be observed by analyzing the dispersion of the emotions.

Finally, Fig. 9 displays the evolution of the dispersion around the group of agents' social emotion. These dispersions are calculated for each component of the PAD vector. It can be observed that they become lower as time passes when the number of people is medium, meaning that this is a very cohesive group that receives in a positive way all the music suggested by the DJ agent. In the other two cases (low and high), there exists a higher dispersion in some of the PAD components. In the first one, in which the number of agents is low, the dispersion levels in each of the components are high except the dominance. A very similar behaviour can be

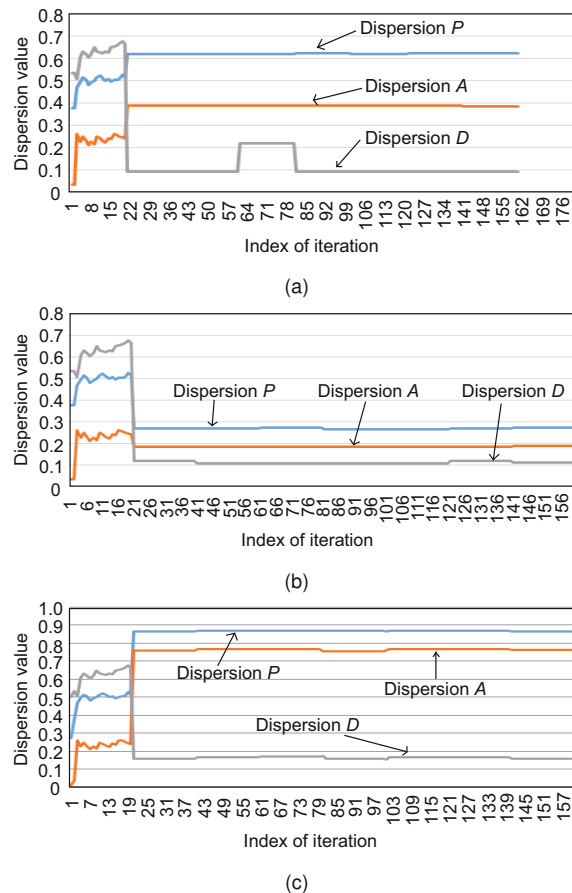


Fig. 9 Experiment 3: evaluation of the dispersion in the emotional state of the group against happiness for each PAD component. The number of people is low ( $\leq 20$ ) (a), medium (20–80) (b), or high ( $> 80$ ) (c)

observed in the scenario where the number of agents is high. This situation can be due by the variability of agents who are comfortable or uncomfortable according to the number of agents placed around them. This can be reflected in that the dispersion in each component of the PAD vector is high.

## 7 Conclusions

Multi-agent systems allow the design and implementation of applications where the main components can be humans and software agents interacting and communicating among them in order to help humans in their daily activities. In this sense, this paper presents an AmI application where humans and agents must coexist in a framework of maximum integration. The application has been developed over the JaCaIIVE framework allowing an easy integration of the human in the multi-agent system and a visualization of the system in a virtual environment. The proposed system is able to extract (in a non-invasive way) and to analyze the social emotion of a group of persons and it can take decisions according to that emotional state. The application has been tested in order to verify how agents improve their decision-making process. In this sense, we have shown how our developed agents change their emotional states according to changes in the environment. This way, these emotional changes are translated into an aggregated view of the emotional state of the group. This social view of the emotions is employed by the system in order to improve actions to be taken in the future. In this case, the improved action is the selection of the more appropriate songs to be played for the group of agents.

Future work in this research area will focus on developing a learning module which will allow the DJ agent to anticipate to future emotional states of the people. This module will improve the decision making of the DJ agent comparing the current situation with similar previous situations.

## References

- Ali, F., Amin, M., 2014. The influence of physical environment on emotions, customer satisfaction and behavioural intentions in Chinese resort hotel industry. *J. Glob. Bus. Adv.*, **7**(3):249-266. <http://dx.doi.org/10.1504/JGBA.2014.064109>
- Bales, R.F., 2001. *Social Interaction Systems: Theory and Measurement*. Transaction Publishers, USA.
- Barella, A., Ricci, A., Boissier, O., et al., 2012. MAM5: multi-agent model for intelligent virtual environments. Proc. 10th European Workshop on Multi-Agent Systems, p.16-30.
- Billhardt, H., Julian, V., Corchado, J.M., et al., 2014. An architecture proposal for human-agent societies. Proc. Int. Workshop on Highlights of Practical Applications of Heterogeneous Multi-Agent Systems, p.344-357. [http://dx.doi.org/10.1007/978-3-319-07767-3\\_31](http://dx.doi.org/10.1007/978-3-319-07767-3_31)
- Ducatel, K., Bogdanowicz, M., Scapolo, F., et al., 2001. Scenarios for Ambient Intelligence in 2010. Office for Official Publications of the European Communities.
- Fernandez, A., Ossowski, S., 2011. A multiagent approach to the dynamic enactment of semantic transportation services. *IEEE Trans. Intell. Transp. Syst.*, **12**(2):333-342. <http://dx.doi.org/10.1109/TITS.2011.2106776>
- Hale, K.S., Stanney, K.M., 2002. *Handbook of Virtual Environments: Design, Implementation, and Applications*. CRC Press, USA.
- Han, D.M., Lim, J.H., 2010. Smart home energy management system using IEEE 802.15.4 and ZigBee. *IEEE Trans. Consum. Electron.*, **56**(3):1403-1410. <http://dx.doi.org/10.1109/TCE.2010.5606276>
- Hendler, J., 2007. Where are all the intelligent agents? *IEEE Intell. Syst.*, **22**:2-3.
- Holzappel, A., Stylianou, Y., 2007. A statistical approach to musical genre classification using non-negative matrix factorization. Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, p.693-696. <http://dx.doi.org/10.1109/ICASSP.2007.366330>
- Huhns, M.N., Singh, M.P., Burstein, M., et al., 2005. Research directions for service-oriented multiagent systems. *IEEE Internet Comput.*, **9**(6):65.
- Intille, S.S., 2002. Designing a home of the future. *IEEE Perva. Comput.*, **1**(2):76-82.
- Lawrence, S., Giles, C.L., Tsoi, A.C., et al., 1997. Face recognition: a convolutional neural-network approach. *IEEE Trans. Neur. Netw.*, **8**(1):98-113. <http://dx.doi.org/10.1109/72.554195>
- Li, T., Ogiwara, M., Li, Q., 2003. A comparative study on content-based music genre classification. Proc. 26th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval, p.282-289. <http://dx.doi.org/10.1145/860435.860487>
- Mangina, E., Carbo, J., Molina, J.M., 2009. *Agent-Based Ubiquitous Computing*. Atlantis Press, France. <http://dx.doi.org/10.2991/978-94-91216-31-2>
- Mehrabian, A., 1980. *Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies*. Oelgeschlager, Gunn & Hain, USA.
- Mehrabian, A., 1997. Analysis of affiliation-related traits in terms of the PAD temperament model. *J. Psychol.*, **131**(1):101-117. <http://dx.doi.org/10.1080/00223989709603508>
- Nanty, A., Gelin, R., 2013. Fuzzy controlled PAD emotional state of a NAO robot. Proc. Conf. on Technologies and Applications of Artificial Intelligence, p.90-96. <http://dx.doi.org/10.1109/TAAL.2013.30>
- Ortony, A., 1990. *The Cognitive Structure of Emotions*. Cambridge University Press, USA.

- Ossowski, S., 2013. Agreement Technologies: 8 (Law, Governance and Technology Series). Springer, the Netherlands. <http://dx.doi.org/10.1007/978-94-007-5583-3>
- Osuna, E., Freund, R., Girosit, F., 1997. Training support vector machines: an application to face detection. Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, p.130-136. <http://dx.doi.org/10.1109/CVPR.1997.609310>
- Rincon, J.A., Garcia, E., Julian, V., et al., 2014. Developing adaptive agents situated in intelligent virtual environments. Proc. 9th Int. Conf. on Hybrid Artificial Intelligence Systems, p.98-109. [http://dx.doi.org/10.1007/978-3-319-07617-1\\_9](http://dx.doi.org/10.1007/978-3-319-07617-1_9)
- Rincon, J.A., Julian, V., Carrascosa, C., 2015a. Applying a social emotional model in human-agent societies. Proc. Int. Workshops of Practical Applications of Agents, Multi-Agent Systems, p.377-388. [http://dx.doi.org/10.1007/978-3-319-19033-4\\_33](http://dx.doi.org/10.1007/978-3-319-19033-4_33)
- Rincon, J.A., Julian, V., Carrascosa, C., 2015b. Social emotional model. Proc. 13th Int. Conf. on Practical Applications of Agents, Multi-Agent Systems, p.199-210. [http://dx.doi.org/10.1007/978-3-319-18944-4\\_17](http://dx.doi.org/10.1007/978-3-319-18944-4_17)
- Satyanarayanan, M., 2001. Pervasive computing: vision and challenges. *IEEE Pers. Commun.*, **8**(4):10-17. <http://dx.doi.org/10.1109/98.943998>
- Satyanarayanan, M., 2002. A catalyst for mobile and ubiquitous computing. *IEEE Perva. Comput.*, **1**(1):2-5. <http://dx.doi.org/10.1109/MPRV.2002.993138>
- Talupur, M., Nath, S., Yan, H., 2001. Classification of Music Genre. Project Report for 15781.
- Viola, P., Jones, M.J., 2004. Robust real-time face detection. *Int. J. Comput. Vis.*, **57**(2):137-154. <http://dx.doi.org/10.1023/B:VISI.0000013087.49260.fb>
- Weiser, M., 1991. The computer for the 21st century. *Sci. Am.*, **265**(3):94-104. <http://dx.doi.org/10.1038/scientificamerican0991-94>
- Zambonelli, F., Jennings, N.R., Wooldridge, M., 2003. Developing multiagent systems: the Gaia methodology. *ACM Trans. Softw. Eng. Meth.*, **12**(3):317-370. <http://dx.doi.org/10.1145/958961.958963>