EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2017)
T. Schreck and T. Weyrich (Editors)

# Animating with a self-organizing population the reconstruction of medieval Mértola

R. F. Antunes[1,2] and A.P. Cláudio[1] and M.B. Carmo[1] and L. Correia[1]

[1]BioISI - University of Lisbon, Portugal
[2]MIRALab, University of Geneva, Switzerland

**Abstract**

*This paper provides a contribution to the field of historical simulations of the past. Throughout this document, we will describe a novel model to animate these simulations with autonomous characters exhibiting heterogeneous and spontaneous behaviours and we will discuss a case study, the simulation of the medieval village of Mértola, in the South of Portugal. We will first detail the work of construction of the urban layout. Using manual modeling combined with procedural generation, we have generated a virtual space containing some of the military structures, such as the defensive walls and the watchtower in the river, as well as some of the civilian housing inside the protection of the walls. Following, we will describe the virtual population inhabiting the space composed of autonomous individuals dressed with historical rigour. These inhabitants of the virtual city, are equipped with limited intelligence and personality traits which allows them to self-organize, interact with each other and at the local market. They communicate with their fellow citizens in the narrow streets of the village using expressive gestures and postures that convey their inner emotional states.*

**CCS Concepts**

•***Computing methodologies*** $\rightarrow$ ***Animation; Procedural animation;*** *Multi-agent systems; Intelligent agents; Artificial life;*

## 1. Introduction

Undertaking a reconstruction of ancient sites in 3D is an endeavor requiring great resolution and venture, in a process typically involving a multidisciplinary team, including 3D modelers, archaeologists, and historians. The benefits compensate the efforts and audiences can get a sense of the environment: the architectural features of the site, the materials and textures used in the construction, the layout of the edifices and their occupation, as well as the surrounding landscape. The integration of virtual humans in 3D simulations of ancient historical sites brings in layers of additional possibilities to these artifacts. Either, *a)* as virtual humans, interacting with the audience, or *b)* acting as background extras, filling in the scene, the presence of virtual humans enhances the artefacts allowing richer and more informative experiences. For instance, virtual humans can be integrated as single virtual guides, interacting with visitors, and providing contextually relevant information about the simulated site [RBKSS09]. With technological progress, the simulation of groups and populations is becoming more accessible everyday. Simulations feature smaller groups of actors [PFMT03] or entire crowds with hundreds of individuals [MHY*07]. The complexity of their actions equally varies from simple actions such as wandering in the environment [RFD05] to differentiated roles in populations of autonomous characters [LCG*13, TBS14]. We aim at contributing to this field with a novel model to animate these simulations with autonomous characters exhibiting a high degree of

**Figure 1:** *Top: Overall perspective of the village. Bottom: Snapshot of the market with population.*



DOI: 10.2312/gch.20171286

heterogeneous and spontaneous behaviors. Throughout this document, we describe the virtual reconstruction of the medieval village of Mértola, in the South of Portugal (Figure 1). We start by reporting on the process of construction of the urban layout of the virtual space, discussing the simulation of some of the civilian housing, together with military structures, such as the defensive walls and the watchtower in the river using manual modeling combined with procedural generation. Then, we will detail the process of animation in real-time of the virtual population of historical characters. This population can be described as a society of self-organizing historical characters, dressed in historical rigor, that interact and trade with each other and at the local market. The model of artificial intelligence described here enables individuals to be autonomous and expressive when communicating with each other. The audience can experience the simulation, either as educational videos, or more interactively, navigating the simulated space in real time, commanding a camera in bird's eye view in a Desktop-based application. The official Facebook page of the project is: `https://www.facebook.com/mcbihc/` , and a video of the simulation is available at: `https://youtu.be/8s9mVzNVH7A`.

## 2. State of the Art

**Inhabited cultural heritage simulations** Historical simulations including autonomous populations in real time are relatively uncommon given the complexity of their implementation. The process of construction usually requires extra work with the additional presence of programmers in teams that already bring together a vast skill-set. Examples of these endeavors include a small group of worshippers entering a mosque and performing their ceremonial rites [PFMT03]. Althought the characters share common gestures, the animations are activated in an asynchronous way and the duration of each sequence is unique for each of the characters. Larger crowds appear in the simulation of the 7th-century village of Wolfenbuttel, in Northen Saxony, with virtual intelligent guides responding in real time to users. The virtual inhabitants wander on the environment, avoiding each other but without any visible interactions [RFD05]. The absence of gregariousness is recurrent and we can see that again, for instance, in a simulation of the Roman Flavian amphitheatre [GF05]. More sophisticated autonomous behaviours are found in virtual Pompeii. The social status restricts the spacial navigation of the agents into poorer or richer areas. Despite that distinction, interactions between agents are undifferentiated regardless of social status or location. A semantic description of the city informs the agent what actions can he perform at specific sites, such as entering a bakery and leaving with bread on the hand [MHY*07]. Semantic descriptions reappear also in the Babylonian city of Uruk in Second Life. Inhabitants are self-motivated and need to perform tasks and communicate [TBS14]. Additionally, they also get hungry and become tired. Reactions to identical stimulus differ as their unique personality influences their choice of actions. However, psychological traits were not fully explored to change expressions, and feeble differentiation is offered at the individual level. This absence of individuality is a recurrent problem, evidenced again in the simulation of the old Pennsylvania station in New York. Characters roam the space performing preconfigured activities and interacting with smart objects, as well as other inhabitants [ST05]. Individuals also have desires that need to be fulfilled. They can buy tickets, buy drinks, watch performances, seat and rest, and for each action, they trigger preconfigured sequences of actions and animations. However, to our knowledge, besides the scripted behaviours, no other spontaneous interactions occur. A compelling atmosphere looms from the simulation of the Port of Georgetown, in Malaysia. Lim and colleagues attempted to simulate the interaction of populations composed of four different classes of characters (ethnical groups). Based on cultural assumptions made by the modellers, each class uses different action state machines. Parameters for navigation and aggregation (flocking behaviour) also differ. Individual parametrization is stochastic within an interval of possible values to avoid further homogenization [LCG*13]. Despite the encouraging results on the global motion, the absence of differentiation on gait and individual expression is overtly apparent. Another common problem with these works is their context-dependency. The AI models have been tailor-made to the specific simulations in which they were implemented. Of course, they can be adapted to new situations, but this requires the presence of a programmer to perform the necessary changes. In summary, state of the art in cultural heritage simulations of ancient sites offers rich environments populated with self-organizing actors that engage in autonomous interactions with the other members of the population and elements of the scenario. However, we can identify a few problems, namely: *a)* the limited individual expression of the characters, *b)* the low level of spontaneous and heterogeneous behaviours, *c)* the fact that they require complex implementations of scripted behaviours, *d)* the context-specificity of the implementations. Taking this landscape into consideration, it seems now pertinent to focus on enhancing the levels of heterogeneity, spontaneity, and individual expression in self-organizing autonomous populations, a goal closely aligned with recent research on group and crowd simulation.

**Psychology and individual agency in group and crowd simulation** The research on generative populations of self-organized individuals, with psychological features and expressive behaviors, has been gaining traction in the last years. Durupinar and colleagues have recently presented a model for emotional representation with psychology and emotional states in the context of groups and crowds [DGAB15]. They successfully generated a mob of protesters with individual expression. The OCEAN model (Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) [CM92] is used to represent individual personalities. In their simulation, emotions are dynamic and, in addition, these authors also integrated the effect of emotional contagion. The implementation was tailor made and somehow complex, with context-based behavior trees to simulate the specificities of the protest scenario. Earlier, Cassenti [Cas09] has explored the implementation of societies of autonomous agents equipped with psychological features and motivated by metabolic requirements. The objective was to simulate military behavior, and for this the emotional state was represented with parameters for joy-distress, hope-fear and like-dislike, resulting from actions performed directly by the agent, or the members of his group. These emotions impact the stress level, which triggers further motivated actions. A similar idea of combining mental representation with biological motivations of hungriness and thirstiness, reappeared in SEstar from Navarro *et al.*. They use stress variables integrated with higher level descriptors of personality [NFM15] to simulate populated urban environments. Antunes and Magnenat-Thalmann integrate affective feedback mechanisms between the psychological states of individuals and the choice of their actions and posterior consequences [AMT16].

Given the above, it may be challenging to *a)* generate context-independent populations, which can be integrated smoothly in a wide range of historical simulations, and *b)* enhance the levels of expressivity and individuality in the populations. In the following section, we will describe the model of AI that we have developed with those goals in mind.

## 3. The reconstitution of the space

During the 9th century, the river Guadiana, was an important commercial route of the Al-Andalous. Mértola and its port occupied a strategic location, at the end of the navigable section, playing an important economic role in the commerce of agricultural and mineral goods. Scarce visual information has reached us from that time, and our goal is to recreate in 3D the village, to try to capture and understand its former dynamic and soul. The architectural structures have been going through a long process of stratification, and those visible suffered from deterioration or various programs of renovation. In addition, this was a society with a culture of non-depiction, and consequently, the only visual information we got access was from posterior periods. We have gathered different sources of information: from archeological excavations, along with information retrieved from archives and historical knowledge on the epoch provided by experts, as well as the help of old sketches from the early 1500s. Departing from this material, we have built a computational 3D model of the space and inhabited it with a virtual population of autonomous inhabitants (Figure 2).

We have produced a virtual space featuring military and non-military architectures, like the protective walls or the watchtower in the river, and civil housing. Manual and procedural modeling was used to generate these achitectural structures. Furthermore, the creation of the virtual space comprised the acquisition of the terrain, the generation of streets and the addition of vegetation. The whole process may be described by a pipeline [CCC*17]. First, the terrain was imported using ArcGIS platform (desktop.arcgis.com) and some of the architectural models were developed with 3D modeling tools like AutoCAD (www.autodesk.com/products/autocad/overview), Blender (www.blender.com) and 3DSMax (www.autodesk.com/products/3ds-max/overview). Afterwards, the terrain and the 3D models were imported into City Engine (www.esri.com/software/cityengine) where the virtual model was completed taking advantage of modeling based on sets of rules. The complete virtual environment was then exported to Unity (unity3d.com), in FBX format. Then it was ready to be populated by a community of autonomous historical characters equipped with our agency model. These components are detailed in the following subsections.

### 3.1. Terrain

The generation of the terrain involved three main steps: import datasets, get an image of the terrain to create a texture and, finally, combine data and image. ArcMap, ArcCatalog and ArcToolBox tools of the ArcGIs platform supported the integration of data imported from different sources. The geographical coordinates of the limits of a square region, containing the village of Mértola, were imported from the Portuguese Entity Direcção Geral do Território (mapas.dgterritorio.pt/Openviewer/cos2007/Guia_de_apoio_utilizacao_WFS_ArcGIS.pdf). Given the geographical coordinates, the corresponding raster dataset was imported from Open Topography (www.opentopography.org), which offers raster datasets with different levels of resolution for digital terrain models.To texture the terrain a high-resolution image, in JPG format, was captured with Google Earth Pro. This image was imported into ArcMap and combined with the terrain data. The produced height map and texture were then imported into City Engine.

### 3.2. Architectural models

The set of 3D architectural models includes houses and military structures. Some of the models, such as the watchtower near the river (Figure 2b) and the houses, were produced in AutoCAD, Blender, and 3DSMax by manual modeling. Some of the houses were modeled with three levels of detail and are rendered with the appropriate level, accordingly to their distance to the camera. For the excavated zone (Figure 2c), we followed information from the archaeological reports about the layout of the houses and their spatial disposition. These houses functioned as templates to the ones filling in the remaining areas. The street layout follows roughly the contemporary arrangement since this is believed to have suffered little change throughout the times [Mac96]. The defensive walls were modeled in City Engine. Only some towers in the Northern section, -which appears in the first plan of Figure 2a- still preserve the shape from the Islamic period. With the help of some old drawings [Mac96, Dar09], these were built using polygonal modeling and functioned as templates for the other ones appearing on the virtual walls. Then, these models of towers were used as assets by City Engine in the code rules to create the defensive walls. As Macias points out on his thesis, the walls have gone through many transformations throughout the times, but the layout has been kept unchanged for the larger part of the perimeter. The Grow Streets tool was used to define the configuration of these walls. This layout was adjusted manually to the contour of the defensive walls. Finally, the resulting streets were used as the basis for the rules that lifted the walls. Figure 2a shows partially the defensive wall with some towers.

## 4. The agency model

We aim to simplify the implementation of virtual populations composed of autonomous individuals. In this section, we discuss a generalistic model for populations of self-organizing characters, generating behavioural heterogeneity and spontaneity with individual expressivity (Figure 3). Members of this society are motivated by their intrinsic need for *a)* socialisation and *b)* exchange of resources. The actions they perform in the world are shaped by their


**Figure 2:** *Areas built based on archaeological evidence. Fig. a) shows the virtual reconstruction with the watchtower and the houses of the Alcaçova surrounded by ellipses. Figs. b) and c) show the current state of these areas. The protective walls preserve, in large part, the contour from the Islamic period.*

unique personality and their emotional states. In addition, when interacting, they take into consideration their moods and the memory of preceding events. The main components of the system are described in detail in the next sections.

### 4.1. Individual

#### 4.1.1. Blueprint

A set of fundamental features defines the members of this population: *a)* length of circadian rhythm impacting metabolic rate and gregariousness, *b)* personality, *c)* type of resources consumed, *d)* type of resources being generated, and *f)* sex *g)* sensitiveness to starvation *h)* sensitiveness to loneliness. A stochastic process initializes the blueprints at the beginning of the simulation.

#### 4.1.2. Navigation

The navigation is structured on Navmesh, an off-the-shelf solution provided by Unity3D. This system implements a technique to represent walkable sections of a 3D environment using polygons, known as Navigation mesh [XS12]. Trajectories are calculated using A* upon the vertices of the navigable mesh. When moving, characters always attempt to reach a physical target. In states of wandering, they choose in a stochastic process one of the objects present in the scene and assume it as their target.

#### 4.1.3. Behaviour

The underlying structure defining the behaviour is a Markov chain (Figure 4 and Table 2), as this provides an efficient method for behaviour control with progressive adaptation. Transitions leaving a state have an assigned probability of being performed. Probabilities are identical at initialization time and change dynamically during runtime. Transitions can become active or inactive at any given moment depending on the satisfaction of specific conditions. For instance, the transition from state *S0-Rest* to state *S1-Move to Partner* only becomes active if the condition 'Buddy exists in the vicinity' is satisfied. Transitions of state can have associated one or multiple conditions from the set listed in Table 1. Complex behaviours, such as an interaction, when performed involve a path through var-

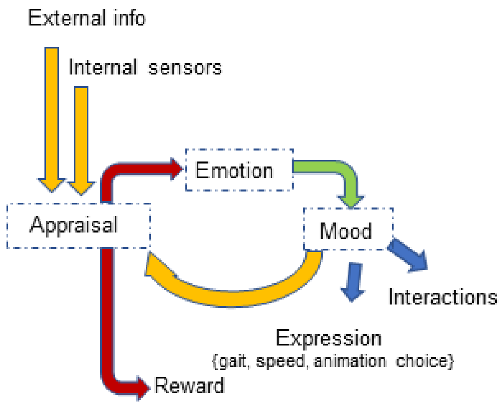


**Figure 3:** *Overall description of the workings of the system. The information of internal and external sensors is considered to produce the appraisal. The appraisal calculates the momentary values for Valence and Arousal. This process generates an emotion and simultaneously, it rewards the transitions from the previous action-state in the Markov chain. The mood is a function of its previous value, plus this new emotion of the individual. The current mood will then be used in future interactions, as well as to give expression to the character.*

**Table 1:** *Rules enabling state-transitions.*

| **Rules conditioning the activation of state-transitions.** |
|---|
| Hungriness < Threshold |
| Gregariousness < Threshold |
| Distance to Target < Threshold |
| Previous knowledge of partner |
| Buddy nearby |
| Trader nearby |

**Table 2:** *States and respective transitions*

| State | Action | Transitions |
|---|---|---|
| 0 | Rest | 0, 1, 2, 3, 4, 8, 12 |
| 1 | Move to partner (social) | 0, 1, 2, 4, 8, 12 |
| 2 | Move to partner (trade) | 0, 1, 2, 4, 8, 12 |
| 3 | Wander | 0, 1, 2, 3, 4, 8, 12 |
| 4 | Found partner (social) | 1, 2, 5, 6, 12 |
| 5 | Interact (social - known partner) | 3, 7, 12 |
| 6 | Interact (social - unknown partner) | 3, 7, 12 |
| 7 | Socialize (chat) | 0, 3, 12 |
| 8 | Found partner (trade) | 1, 2, 4, 9, 10, 12 |
| 9 | Interact (trade – known partner) | 3, 11, 12 |
| 10 | Interact (trade – unknown partner) | 3, 11, 12 |
| 11 | Trade | 0, 3, 12 |
| 12 | Emergency | 0, 1, 2 |

ious states of the Markov chain. For instance, when trading, the interaction unfolds in three moments/states: *i)* encounter of the two individuals, *ii)* decision if they previously met and if they desire to trade, *iii)* exchange of resources/ abandon of interaction.

Once an action-state is performed, the transitions from the previous state are rewarded per the affective value of the action performed. A positive pay is attributed to the link responsible for transitioning to the new state, while the other remaining links are negatively paid. Reward of a link $l$ transitioning to state $s$ is given by Equation 1,

$$R(l,s) = \theta^{v} v(s) + \theta^{a} a(s) R \rightarrow [0,1]. \tag{1}$$

where $\theta^{v}$ and $\theta^{a}$ are scalar weight coefficients. v(s) (Equation 2) and a(s) (Equation 3), are functions measuring the affect that results from performing the action associated with state $s$.

#### 4.1.4. Psychology

Emotional dynamics provides layers of individual expression to the characters and their behaviors. To guarantee smooth transitions, the psychological state is designed as a bipartite scheme composed of two components, one steadier, the mood, and the other volatile, the emotion. To assure the audience faces coherent behaviours and expressions, and provide continuity between the current emotion and the previous ones, both dimensions are interdependent. The psychology draws upon Russel Circumplex model [Rus80]. Under Russel's model, emotions are distributed in a two-dimensional circular space. The two axis represent dimensions for Arousal (a) and Valence (v). The vertical-axis represents a continuum between low and high arousal. The horizontal axis offers a continuum for representing pleasantness. Our system represents these dimensions in the space [0,1], and the emotional range spans from euphoria (v=1, a=1) to deep depression (v=0, a=0).

**Valence** The extent to which an emotion is positive or negative is described as emotional valence [FR99]. Equation 2 defines this dimension in our model. Interactions play a key role in the workings of this architecture as their outcome not only is used in the reward process of the action-state performed, as it also impacts the emotion of the agent. We distinguish three contexts. The first is when the individual is interacting, trading resources. In this case, valence is a function of the total of resources effectively exchanged, $k$. To induce a measure of pleasantness to transactions, we have tailored the equation to converge to zero the lesser the traded units are. In turn, when in the context of socialisation, valence is a function of the gregariousness and the loneliness of the individual. The equation converges to one when the individual fits within a comfortable zone (gregariousness level above the threshold urgency). In equation 2, $g$ is his current gregariousness level and $w$ the bonding-need threshold, below which the individual feels socially isolated and lonely. $\lambda$ is a binary indicator of dialogue with value {0,1}. Finally, when the individual is in a state of walking, valence is a function of the distance to his target $d$, with the function generating growing values of valence the closer the individual is relative to its target.

$$v(s) = \begin{cases} \begin{cases} \left(1-\frac{1}{k+1}\right)0.75 & ,k>0 \\ 0 & ,k=0 \end{cases} & \text{, trading} \\ \lambda(1-\frac{1}{10}^{\frac{g}{w}}) & \text{, socialising} \\ \frac{1}{d} & \text{, otherwise} \end{cases} \quad v \rightarrow [0,1] \tag{2}$$

**Arousal** The second appraisal equation calculates the arousal of the emotion. This dimension measures the intensity of an emotion and is given by Equation 3. As before, we also distinguish three contexts, trading, socialisation and walking. When trading or socialising, the arousal converges to zero the more comfortable the agent gets. Comfort is measured in terms of the current levels of resources and gregariousness. To implement this effect, the level of comfort acts as the exponent of the fractions. In either case (trade/socialize), we multiply the fraction by two, in order to generate values in the positive quadrant (greater than 0.5 in the space [0,1]) when the level of energy/gregariousness is greater than their respective thresholds. When walking, we take both aspects of motivation into consideration. In the function, $e$ stands for the current energetic level, and $z$ for the threshold level below which the agent feels "hungry". $g$ and $w$ stand for gregariousness level and bonding-need threshold. Values produced by the function are capped to one.

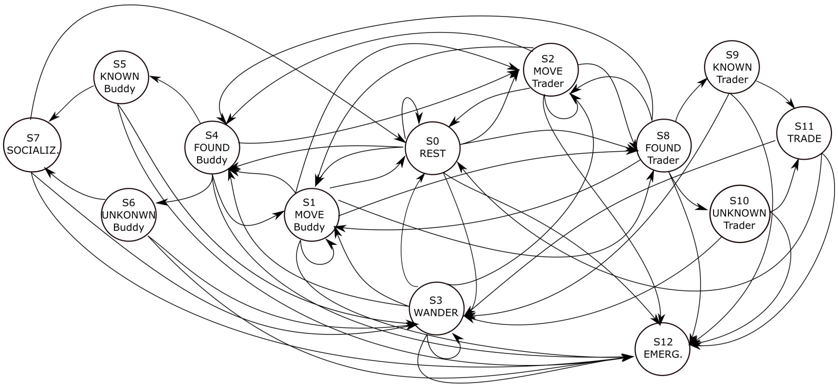


**Figure 4:** *Diagram of the Markov chain, showing the action-states and their respective transitions.*

$$a(s) = \begin{cases} 2\left(\frac{1}{4}^{\frac{e}{z}}\right), \text{trading} \\ 2\left(\frac{1}{4}^{\frac{g}{w}}\right), \text{socialising} \\ \left(\frac{1}{4}^{\frac{e}{z}}\right) + \left(\frac{1}{4}^{\frac{g}{w}}\right), \text{otherwise} \end{cases} \qquad a \rightarrow [0,1] \tag{3}$$

**Emotions and Mood** Each of the psychological dimensions is represented using bi-dimensional vectors: $\vec{M}$, for mood (Equation 5) and $\vec{E}$, for emotion ( Equation 4).

$$\overrightarrow{E}_t = \gamma\overrightarrow{v(s), a(s)} + \delta\overrightarrow{M}_{t-1}\colon \mathrm{E} \rightarrow [0,1]. \tag{4}$$

where $\gamma$ and $\delta$ are weight coefficients.

We have opted for modelling mood as a quality that is initially inscribed in the genetic blueprint of the individual. However, its nature is dynamic and slowly changes in time as a result of the cumulative effect of posterior emotions. In our system, a mood can be described as a temperamental factor that slowly reflects the memory of previous emotions. At time $t$ the mood is a function of its previous value $\vec{M}_{(t-1)}$ and the emotion $\vec{E}_{t-1}$. In Equation 5, $\alpha$ and $\beta$ are weight coefficients.

$$\overrightarrow{M}_t = \alpha\overrightarrow{M}_{(t-1)} + \beta\overrightarrow{E}_{t-1}\colon \mathrm{M} \rightarrow [0,1]. \tag{5}$$

Mood is one cornerstone structure of this architecture. It plays a critical role in giving expression to the character. The attributes of locomotion, gait, and speed, mirror this psychological dimension of the individual (Figure 7). The mood impacts, as well, with the choice of gestural animation being used during the interactions. We will discuss this in Section 4.3.

#### 4.1.5. Metabolism and economy of resources

Each action performed by the individual in the world has an associated energetic cost. The agents gather specific resources and later convert these into useful energy in an operation that also generates a secondary type of resource. There are three different type of resources {a,b,c} and each individual consumes and stores two of these, while he produces the third type. For instance, to perform the metabolic function, an arbitrary individual uses one unit of resource type *a*, and one unit of resource type *b* to produce one unit of resource type *c*, and ten units of *energy*. The individual can collect new token-units when trading with other citizens exchanging the ones he owns by the ones he needs. Alternatively, some places simulate working activity for agents. Here, agents also obtain token-units, but at the expense of a steep energetic price.

### 4.2. The society

The society is composed of two types of individuals: Producers and Citizens. Producers are animated as merchants or farmers, and functionally they act as fountains of resources producing a constant amount of token-units at periodic intervals of time (Figure 5 - Left). Their animation states are predefined using simple automata with three sequential states. Citizens on the contrary, are dynamic and members of the regular population (Figure 5 - Right). These are the ones with the agency model implemented and show emotional dynamics.

**Figure 5:** Left- *A merchant;* Right- *A citizen.*

#### 4.2.1. Interactions

We emulate some of the fundamental human motivations. The basic premises of this model are: *a)* that individuals use resources and convert them into useful energy, and *b)* that they are gregarious, socialising from time to time. These features represent some of the primal impulses, prioritary in the terms defined in Maslow pyramid of needs [Mas43], the physiological necessity for nourishing, and the psychological need for a sense of belongingness [BL95]. To model these essential aspects, we have individuals driven by variables of energy and gregariousness, and the nature of their motivations to interact can be one of the following: *a)* socialisation and *b)* trading.

**Trading** Agents need resources to convert into useful energy. An individual takes initiative to trade when his energy levels fall below a threshold *t0*. Any individual can be approached at any time by some other motivated member of the population. When trading, the individual exchanges the resources he needs by the ones he has in excess. However, he will only agree to trade if he needs some of the resources provided by the other. During the transaction, for each type of resource, there are two other important thresholds: t1 (thereshold-of-need), t2 (threshold-of-surplus). If both agents can simultaneously receive needed resources and provide useful ones, then the transaction progresses. Otherwise it fails. A new emotion is generated from this encounter, with the value of valence being pushed to the negative quadrants (below 0.5 in the space [0,1]) if the transaction fails, or positive when there are resources being transactioned, with the value of valence being relative to the total of resources effectively transactioned (Equation 4 and Equation 2).

**Socialisation** When agents $i$ and $j$ interact, this was most likely caused by $i$ taking initiative given his bonding levels falling below an energetic threshold *t1* (bonding-urgency). As with the trading interaction, any agent can be approached at any time. When this is the case with agent $j$, he only accepts to enter on dialogue if his threshold drops down of an excess cap *t2* (excessive-bonding). The emotion generated by this encounter is relative to the solitude and isolation felt by each of the individuals (Equation 4).

### 4.3. Animations

Each state in the Markov chain corresponds also to an animation-state (Figure 6). Each of these animation-states branches on five different animation clips which selection is dependent on either the valence or the arousal parameter of the current mood. The mood is the corner-stone of this design, as it provides expression to the character. The dimension of arousal is taken into consideration for functions of locomotion, and the dimension of valence for interactions. The animator mechanism of Unity3D (Mecanim) allows the creation of interpolations between animation parameters. Consequently, the input of the vector is interpreted in a continuous space, since the animations will blend in the interpolated regions between clip definition. For instance, if animation A is assigned to value zero of arousal, and an animation B to value one, then when arousal has a value of 0.5 Mecanim will interpolate the parameters of the two animations, and blend them in average values. Figure 7 shows how the gait of the individual mirrors this psychological dimension. Furthermore, when the current active state is an interaction state, the animation sequence subdivides per different interaction conditions: *a)* which of the partners took the initiative, and *b)* whether the agent decides to establish a dialogue or not. As above, each of these branches also provides five emotionally related animations.

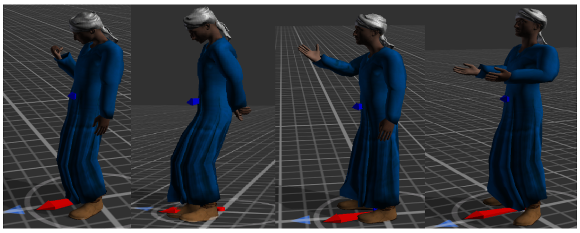

**Figure 6:** *Four examples of interaction clips that are triggered according to the animation-state activated by the valence component of the current mood.*

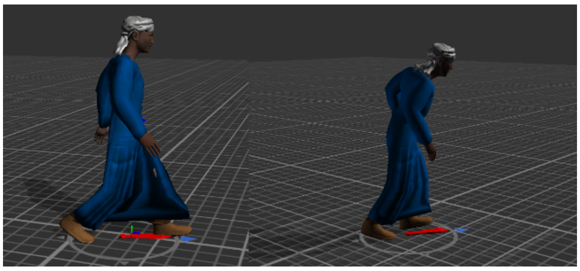

**Figure 7:** *Two images demonstrate the impact of arousal on the gait of one character. On the left, the straight posture when the individual has high positive arousal and on the right, the same character with a negative value of arousal.*

## 5. Results

To study the behavior of the model, we have set a simulation with a population of fifty individuals. We took snapshots of the internal parameters for each of the individuals at intervals of thirty seconds, and we set the simulation running for a period of one hour in order to analyze the outcome. For this experimental setup we used the following parameters: $\theta^v = 0.5$, $\theta^a = 0.5$, $\gamma = 0.55$ and $\delta = 0.45$. $\alpha = 0.95$ and $\beta$=0.05. The dataset and additional materials can be found at `https://doi.org/10.5281/zenodo.832720`.

The first aspect we went to look closer was the progression of emotions. Do individuals jump from one emotion to its opposite; or is there a continuous and smooth emotional progression? Figure 8 offers a portrait of the psychological space of two arbitrary individuals during the run. This picture depicts the sequential evolution of their mood and emotions. As expected, emotions are more energetic, while moods also show dynamism but flow in a more composed way. We can see clearly, that rooted on their moods, their emotional repertoire becomes anchored in distinct zones of the circumplex domain, providing peculiar personalities to the two individuals.

We also have looked closer at the relationship of actions performed in the world and their emotional responses. Figure 9 portraits the internal states of an individual that is initially walking, and while performing such activity little changes in his psychological state. Then, he engages in a trading interaction and we see a jump in the value of valence, characteristic from a successful transaction. The mood only suffers a slight positive disturbance as

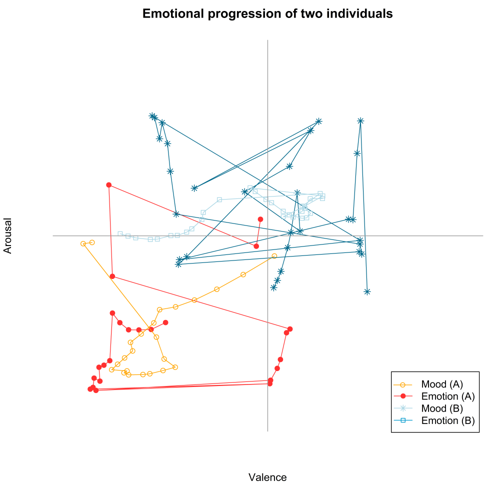


**Figure 8:** *Portrait of the emotional state of two arbitrary individuals during the run.*

a consequence. Following, valence drops in two unsuccessful attempts to trade, to raise sharply when it becomes successfull. Then, it raises again with an acknowledgement to a new partner, just to fall with the consequent unsuccess. It is interesting to notice that the arousal dimension acted as a mirror reflecting the inverse of the bonding level. Every time the agent interacted, receiving a bonding boost, the arousal dropped, and the inverse happened with the lowering of bonding. All this time, the mood has been echoing these changes, dimly reflecting them in its values. This figure shows a dynamic feedback mechanism between internal motivations, actions, and emotions, which is one of the core ideas informing our model.

## 6. Discussion

As mentioned earlier, adding autonomous populations is an important feature for cultural heritage simulations. The advantages are clear, and we can avoid the ghost-city look of many previous historical simulations providing a rich and varied atmosphere of a living population. The audiences can learn about the demographics, the dressing codes, have a clear perception of the occupation of spaces, and even become aware of some of the occupations of the inhabitants. Simultaneously, these simulations can promote the engagement with cultural heritage among new generations highly familiar with game-based technologies. In this document, we have presented an AI model that integrates aspects of metabolism and emotional representation to produce general purpose generative populations showing expressive behaviors and being composed of autonomous and self-organized individuals. The real-time and generative self-organizing nature of the model is meaningful in the sense that two distinct runs will offer differentiated spatial distribution and collective arrangements, boosting the appeal of the 3D reconstructions and prickling the curiosity and interest in the audiences. This type of solution is flexible enough allowing different format outcomes, such as software application or the production of educational videos. These products can be exhibited *in situ* at the museum or real sites in projections, interactive displays, or immersive VR experiences, or yet, be distributed via network or CD to be accessed in the comfort of the home-desktop. The virtues of adding individuality and a layer of expression to the gait become apparent in a video we made as a benchmark, contrasting our model against a method only focused on locomotion. This is available at: `http://xelb.campus.ciencias.ulisboa.pt/videos/benchmark.mp4`. Both groups of characters are called simultaneously to targets located behind the camera. On the left side, we can see the RVO characters walking unnaturally in a rhythmic and almost robotic way. On the contrary, on the right, we can see one second class of individuals animated with an implementation of our model, far more spontaneous and naturally looking, showing differentiation in gait and expression, with emotional continuity, and flow. Although the proposed model has some limitations, namely in terms of group interaction, such as walking in pairs, it provides autonomous agents at a reasonable complexity-behavior variety trade-off. Actions are simple enough to make the model easily adaptable, and the context-independency offered by this generalistic solution offers rich possibilities to those willing to bring to life this type of simulation. The model is easily replicable, and in contrast with the previous examples discussed in the state-of-art, our contribution is somehow generalist to the extent that can be integrated with varied environments, with little or no adaptation. Our model provides a ground layer of animation upon which other behaviors can be integrated. The system is flexible and modular enough so that new characters, with behaviour repertoires designed for specific contexts, can be added and run. Resources defined in Section 4 are symbolic, and it is up to the modelers of future simulations to define what type of producers will they include. With that premise in mind, we tried to offer generalistic behaviors such as conversations as these seem to be actions widespread among human societies. However, gestures in the simulation can be at times effusive. This fact raises questions, given that it might induce misconceptions about the social conduct in the represented societies. The intensity of gestures is culture specific. Social behavior is not universal, and it is a constitutive aspect of the immaterial cultural heritage. Other example being male-female public interactions which were/are tabu in many traditional societies. The importance of gaze in defining socio-hierarchical relationships is another critical aspect. A vast amount of literature exists on culture and social interactions (e.g [Hal73]), and some modest attempts have been made to integrate this knowledge in crowd simulation (e.g [GKLM11]). Regarding cultural heritage simulations, there is need to consider beyond reconstruction of space, a reconstruction of what is going on at the social level. There is room for improvement, and a door is open for future fundamental work on the parametrization and creation of rules that allow non-programmer users to model the ideosyncratic nature of populations. Recent work on crowd research seems to point in this direction with the impact of culture on the human motion being analyzed from real footage in different geographic locations [FKZ13], or the synthesis of motion based on stereotypes on nationality and profession [DWNB16]. It is of critical importance to address these issues in future work. As a proof-of-concept to our model, we have managed to reconstruct a small historical village (of about 2000 to 2500 inhabitants at its peak, according to Santiago Macias [Mac96]). Our final 3D simulation has about two hundred houses and the protective walls with one hun-

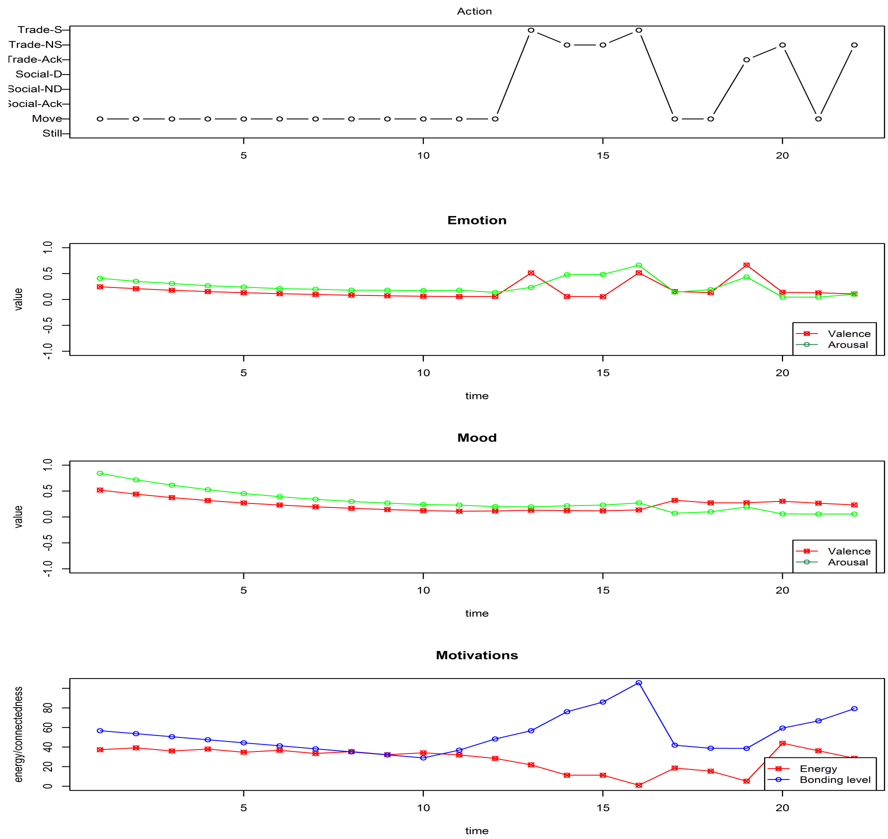


**Figure 9:** *Portrait of the emotional state of one arbitrary individual and the activities he is performing at any moment in time. At the top picture, we can identify the activity being performed. From top to bottom: succefull trade, traiding failure, acknowledgment before trading, diagogue, failed dialogue, acknowledgment before talk, movement, and still. The two following graphs, depict the momentary emotions, and mood, respectively. In these graphs the colors red and green represent the dimensions for Valence and Arousal. The graph in the bottom depicts the energetic and gregariousness levels, which are the main triggers for motivation.*

dred simultaneous inhabitants active inside the wall, dressed in historical rigor (the data-set discussed here only has fifty individuals for the convenience of representation). As Macias points out, primary activities were prevalent. Shepherdess, cattle ranching, agriculture, and fishing were an essential part of the life of the local community. Trade, too, occupied a relevant sector of the community. Besides the utilitarian goods such as baskets or pottery, he reminds us that street selling of food was a transversal practice in medieval societies, and in the case of Mértola, there was even specific regulation for this activity. In the simulation, we have included producers of the type merchants, fisherman, and farmers. Only a few were added to illustrate these occupations since we could not find evidence for the ratio of these among the population. This comment from Macias also raises the question of non-human life in simulations of ancient societies. We know that animal life is a central part of daily life, and to our knowledge, little attention has been paid to this aspect so far. There is still great progress to be made on this type of work. This research is work-in-progress and, despite the fact that the current model offers relevant and useful features for social-historical simulation, we still need to make a user-study with a population of historians and archaeologists that can directly benefit from this type of tools while enhancing and informing future developments.

## 7. Conclusions

This paper presents a method to simulate life in 3D reconstructions of cultural heritage sites. We propose a new model which creates context-free populations that can be easily implemented regardless of their historical context. We do not claim this to be an accurate model of human cognition. Rather, we were focused on generating natural looking human-like atmospheres limited by restraints on real-time computations of animations. This generalistic method follows a bottom-up approach where the livelihood of the scene emerges from the aggregate of motion resulting from the actions of inhabitants. Characters are autonomous, self-organized and expressive. They are equipped with psychological traits, operating in feedback loops between actions and emotions. Two basic motivations drive the actions of the individuals in this society, their need for resources -to convert into useful energy-, and their need for social contact. We further have described an animated population to illustrate ancient life in medieval Mértola. The human presence on these type of simulations raises several questions. There is a justified reluctance of archaeologists in supporting subjects they have no evidence for, such as human behaviour and interaction in the past. However, the lack of human presence in simulations of ancient sites depreciates the overall realism. Socially plausible behaviours can bring the illusion of presence avoiding the ghost city look. Making them more lively is an important task not only to generate the general atmosphere of the place but more importantly to test the potential hypothesis about the use of the sites in ancient times. This sort of exploration requires research contributions that consider historical sources and surveys on social codes of dressing and behaviour in ancient societies. Our model is relatively simplistic regarding sociological realism. To make it more useful we would need to integrate agents that walk in pairs and groups, are capable of a broad variety of activities , and are sensitive to cultural differences such as variations in tolerance to proximity, body contact, gaze, and gestures. Nevertheless, we have managed to generate expressive characters equipped with a bipartite model of psychology, including moods and emotion. Moods functions as an anchor, rooting the emotional tendency. Next steps will also include modelling facial expression, a feature more urgent as immersive VR becomes more pervasive and widespread and the boundaries between the micro and macro levels of the multitude become blurred. Simulations need to accommodate different points of view during one single runtime, for instance when the viewer is immersed in the multitude or in on the top of a tower.

## 8. Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 655226. We also would like to thank the anonymous reviewers for their constructive feedback.

## References

[AMT16] Antunes R. F., Magnenat-Thalmann N.: Bio-inspired virtual populations: Adaptive behavior with affective feedback. In *Proc. of the 29th Inter. Conf. on Computer Animation and Social Agents (CASA 2016)* (2016), ACM, pp. 101–110. 3

[BL95] Baumeister R., Leary M. R.: The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin 117*, 3 (1995), 1497–529. 7

[Cas09] Cassenti D.: Performance moderated functions server (pmf-serv) military utility: A model and discussion. *Army Research Laborator* (2009). 2

[CCC*17] Cláudio A., Carmo M., Carvalho A., Xavier W., Antunes R.: Recreating a medieval urban scene with virtual intelligent characters: steps to create the complete scenario. *Virtual Archaeology Review 8(*, 17 (2017), 31–41. 3

[CM92] Costa P., McCrae R.: Revised neo personality inventory (neo pi-r) and neo five-factor inventory (neo ffi). *Psychological Assessment Resources* (1992). 2

[Dar09] Darmas D.: *Livro das Fortalezas Situadas no Extremo de Portugal e Castela por Duarte de Armas, Escudeiro da Casa do Rei D. Manuel I*. http://http://digitarq.arquivos.pt/viewer?id=3909707, 1509. 3

[DGAB15] Durupinar F., Gudukbay U., Aman A., Badler N.: Psychological parameters for crowd simulation: From audiences to mobs. In *IEEE Transactions on Visualization and Computer Graphics* (2015). 2

[DWNB16] Durupinar F., Wang K., Nenkova A., Badler N.: An environment for transforming game character animations based on nacionality and profession personality stereotypes. In *Proc. of the 12th AAAI Conf. on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-16)* (2016). 8

[FKZ13] Fridman N., Kaminka G. A., Zilka A.: The impact of culture on crowd dynamics: An empirical approach. In *Proc. of the 2013 Inter. Conf. on Autonomous Agents and Multi-agent Systems* (2013). 8

[FR99] Feldman B., Russell J.: The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science 8* (1999), 10–14. 5

[GF05] Gutierrez G., Fricher B.: Virtual crowds in a digital colosseum. *Virtual Retrospect 82*, 2005 (2005), 87. 2

[GKLM11] Guy S., Kim S., Lin M., Manocha D.: Simulating heterogeneous crowd behaviors using personality trait theory. In *Proc. of the 2011 ACM SIGGRAPH/Eurographics* (2011). 8

[Hal73] Hall E. T.: *The Silent Language*. Anchor Books, 1973. 8

[LCG*13] Lim D., Cani M., Galvane Q., Pettre J., Talib A.: Simulation of past life: controlling agent behaviors from the interaction between ethnic groups. In *Proc. of Digital Heritage Cong.* (2013). 1, 2

[Mac96] Macias S.: *Mértola Islâmica: Estudo Histórico-Arqueológico do Bairro da Alcáçova, Séculos XII-XIII ARQUEOLÓGICA 2.0*. Campo Arqueológico de Mértola, 1996. 3, 8

[Mas43] Maslow A.: A theory of human motivation. *A Psychological review 50*, 4 (1943), 370–396. 7

[MHY*07] Maïm J., Haegler S., Yersin B., Mueller P., Thalmann D., Gool L.: Populating ancient pompeii with crowds of virtual romans. In *Proc. of the 8th Inter. Symp. on Virtual Reality, Archeology and Cultural Heritage (VAST)*. (2007). 1, 2

[NFM15] Navarro L., Flacher F., Meyer C.: Se-star: A large-scale human behavior simulation for planning, decision-making and training. In *Proc. of the 14th Inter. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2015)* (2015), pp. 1939–194. 3

[PFMT03] Papagiannakis G., Foni A., Magnenat-Thalmann N.: Real-time recreated ceremonies in vr reconstituted cultural heritage sites. In *Proc. of CIPA 03* (2003). 1, 2

[RBKSS09] Ramic-Brkic B., Karkin Z., Sadzak A., Selimovic D.and Rizvic S.: Augmented real-time virtual environment of the church of the holy trinity in mostar. In *The 10th Inter. Symp. on Virtual Reality, Archaeology and Cultural Heritage VAST* (2009), K. D., C. P., Pitzalis D., Spina S., (Eds.). 1

[RFD05] Ryder G., Flack P., Day A.: A framework for real-time virtual crowds in cultural heritage environments. In *Proc. of the 6th Inter. Symp. on Virtual Reality, Archaeology and Cultural Heritage (VAST 05)*. (2005). 1, 2

[Rus80] Russell J.: A circumplex model of affect. *Journal of Personality and Social Psychology 39* (1980), 1161–1178. 5

[ST05] Shao W., Terzopoulos D.: Autonomous pedestrians. In *Proc. of the 2005 ACM SIGGRAPH/Eurographics symp. on Computer animation* (2005). 2

[TBS14] Trescak T., Bogdanovych A., Simoff S.: City of uruk 3000 b.c.: Using genetic algorithms, dynamic planning and crowd simulation to re-enact everyday life of ancient sumerians. In *Proc. of the Simulating the Past to Understand Human History Conf. (SPUH 2014)*. (2014). 1, 2

[XS12] Xiao C., Shi H.: An overview of pathfinding in navigation mesh. *International Journal of Computer Science and Network Security 2*, 12 (2012), 48–51. 4