

A Level of Interaction Framework for Exploratory Learning with Characters in Virtual Environments

David Panzoli, Christopher Peters, Ian Dunwell,
Stéphane Sanchez, Panagiotis Petridis, Aristidis Protopsaltis,
Vincent Scesa, and Sara de Freitas

Abstract. This paper investigates a range of challenges faced in the design of a serious game aimed at teaching history *in situ*, through the use of an immersive, open virtual environment. In the context of this paper, such an environment is described as an exploratory, expansive virtual world within which a user may interact in a non-linear and situated fashion with the virtual characters that populate it. The main contribution of this paper consists of the introduction of the Levels of Interaction (LoI) framework, designed to assist in the creation of multiple forms of interaction between a user-driven avatar and synthetic characters. The LoI approach addresses the necessity for balancing computational efficiency with the need to provide believable and interactive virtual characters, allowing varying degrees of visual, interactive and behavioural fidelity. The *Roma Nova* project demonstrates a first implementation of the concept, showing in practice how the LoI are likely to foster more natural interactions between the player and the non-playing characters.

D. Panzoli · C. Peters · I. Dunwell · P. Petridis ·
A. Protopsaltis · S. de Freitas
Serious Games Institute,
Coventry University Technology Park, Cheetah Road, Coventry,
CV1 2TL, United Kingdom
e-mail: DPanzoli@cad.coventry.ac.uk, Sara.deFreitas@coventry.ac.uk

Stéphane Sanchez
IRIT-CNRS, UMR5505, Université de Toulouse,
2 rue du Doyen Gabriel Marty, 31042, Toulouse, France
e-mail: Stephane.Sanchez@univ-tlse1.fr

Vincent Scesa
Instinct[Maker], Toulouse, France
e-mail: v.scesa@instinct-maker.com

1 Introduction

The *Roma Nova* project aims at teaching history to 11 to 14 years old children, which corresponds to key stage 3 (KS3) of the English curriculum. *Roma Nova* is a serious game which enables exploratory learning [10] by immersing the learner/player inside a virtual environment where they learn different aspects of history through their interactions with a crowd of virtual yet authentic Roman characters. It is a single-player game, presented in English language.

The game is built upon *Rome Reborn* [12], the most high-fidelity model of Ancient Rome currently in existence, providing a 3D digital model which may be explored in real-time. *Rome Reborn* includes hundreds of buildings, 32 of which are highly detailed monuments reconstructed on the basis of reliable archaeological evidence, for example, as depicted in figure 1.a. The rest of the 25 to 30 square kilometers model is filled with procedurally-generated buildings based on accurate historical knowledge (figure 1.b).



Fig. 1 The Rome Reborn project [36] is a 3D model of the A.D. 320 city of Rome. The model includes accurate models of monuments (a) as well as procedurally-generated areas (b).

The *Roma Nova* project addresses several challenges faced when seeking to design a serious game where the player is taught history by interacting with autonomous characters in a wide open environment. It builds on previous work at Coventry University [1] that demonstrates the potential of game techniques for cultural heritage experiences, outlining the problems encountered when integrating a substantial number of different state-of-the-art techniques. In [24], we have already suggested a step towards achieving this integration involves the design of a general model, to frame the interactions between the player and the characters, bringing together these different interaction modalities into a single multi-scale agent model.

The first critical aspect of the *Roma Nova* project considers the design of an intelligent tutoring system to explore issues related to pedagogic design, such as how desired learning outcomes may be broken down into different scenarios, how different classes of characters or personality may help learners explore different perspectives on events and ultimately how to assess the knowledge effectively learnt. The present paper focuses on the technical aspects of the game (see [25] for the pedagogical aspects), notably which techniques from behavioural animation or artificial life may be used to enable such a broad range of interactions with any character of the crowd, and how they can be integrated together into a single real-time agent controller, while still enabling many characters to be displayed simultaneously.

2 Related Work

A significant volume of literature exists around the study of projects aiming to populate a virtual environment with a crowd of characters. Crowds are desirable for a range of purposes, such as bringing immersion and presence to a historical place, accurately simulating the behaviour of human people, or application-driven scenarios, such as emergency evacuations [35]. The following projects are particularly worthy of mention since they focus upon simulating realistic and believable crowds.

The Pompeii project [21] “revives” the ancient city of Pompeii, buried during an eruption of the volcano Mount Vesuvius. One of the challenges of the project was to populate the 3D reconstruction of the city with a realistic crowd of Romans. To that end, the project places a particular attention on the diversity of the crowd (figure 2.a and .b). The virtual Romans are modelled using a variety of body, face and clothing models, in such a way that every character appears unique. The project also includes a differentiation between city areas and the subsequent impact on appearances, e.g. the variation between wealthy and poorer areas. Another interesting feature includes the implementation of “places of interest”, wherein certain salient city features can attract the attention of passers-by.

The Pennsylvania station project [34] aims to populate the historical reconstruction of the famous New York city railroad station by the same name. Although an emphasis is placed on the characters’ diversity, akin to the Pompeii project, Shao and Terzopoulos predominantly focus upon emulating the rich variety in behaviour of the characters (figure 2.c). The characters walk toward goals in the station based on individual perceptions of it. They can rest, sit on public benches, talk with each other, queue in front of ticket desks, and so on, so that they appear to have a purpose and intent. To further accentuate the realism, live events take place randomly, such as attracting the attention of passing characters.

Other projects consider the role of the human viewer in more detail when synthesising virtual crowds. The Metropolis project [22] investigates visual



Fig. 2 (a-b) The Pompeii project aims to populate a virtual replica of the buried city of Pompeii, Italy. More than 4,000 virtual Romans populate the city, each unique in their combination of shape, size and colour. (c) The Pennsylvania station is populated with autonomous pedestrians that show different behaviours like queuing, resting, talking, etc. (d) The Metropolis project pays a particular attention to the role of user visual and auditory perception when interacting within the environment.

and auditory perception of crowds in city environments (figure 2.d). It introduces an important methodology for evaluating crowd scenes based on viewer perceptions of realism. For example, at far distances from the camera, it may not be necessary to animate or display models at the same degree of detail as when closer, since the viewer may not be capable of noticing differences. Another aspect of realism investigated is viewer perception assessed with regard to the formation of groups [28].

Although the previous projects explore interesting aspects of a crowd of characters, they do not consider interactions between the player and the crowd. Commercial games are thus worthy of interest, since interactions between the user and automated ‘non-player’ characters (NPCs) have a critical impact on the game-play and the overall experience of the player. Owing

to significant progress in computer hardware over the few last years, recent games allow the player to wander freely in large virtual cities or populated areas. Although such games vary in nature and context, the approach used to manage the interactions between the player and the NPCs is often similar. Traditionally, the player is not allowed to interact with every character in the game. There is a distinction between characters that have been placed at a specific location for interaction and other characters that are simply bringing more life to the area.

A crowd of characters may enhance the perception of a coherent world and the player's feeling of being part of it can even be enhanced if the crowd is able to react to them. The following games introduce interesting crowd behaviours. *Assassin's Creed*, from Ubisoft focus on creating an immersive experience for game players. The crowd of characters the player is navigating through can be considered as a fully-fledged character in the game. Indeed, although the NPCs in the crowd attend their own business, regardless of the player, they are still likely to mind the player, and protest when the player collides with them. Similarly in *Grand Theft Auto*, from Rockstar Games, passing NPCs can adopt a specific behaviour depending on the player's actions or the surrounding atmosphere. For instance, they run away shouting when a gunfight takes place or defend themselves when attacked by the player. *Dead Rising*, from Capcom, uses the crowd as an element of game-play. The crowd of zombie NPCs flocks to converge on the player or circulate around a large area.

Different means are used to signify to the player the characters' ability for engaging in more detailed interaction. In *Fable*, from Lionhead studios, halos are used to highlight such characters (figure 3.a). Additionally, when the player enters a range of proximity to the character, a sign appears above them to specify the kind of interaction expected. Although Blizzard's *World of Warcraft* is a massive multi-player game, mainly populated with human



Fig. 3 (a) In the game *Fable*, the interaction abilities are represented by a coloured glowing halo around the NPC. (b) In *World of Warcraft*, an exclamation mark is used similarly.

players, exclamation marks are used to identify non-playing characters responsible for allocating quests or trading with the player (figure 3.b). This game also demonstrates that, although human interaction may be a predominant component of the game design, a need persists for artificially-driven characters to perform mundane or unrewarding tasks. Characters performing these tasks are central to creating a believable, immersive world; however, these games are typically designed to empower the player as a hero, and to achieve this goal for all human players, the artificial intelligence must serve as the ‘ordinary’ people whom players rise above, rescue or defeat. Therefore, even in massively multiplayer scenarios, a need persists for interactive, believable non-player characters.

In both exploratory and narrative-driven serious games, interaction between the player and artificial non-player characters can form a fundamental part of both game-play and instructional design. A substantial volume of research exists around the theme of populating virtual environments with perceptually and functionally realistic characters - the crowd modelling techniques described in Section 3.1, for example, commonly seek to implement realistic crowd behaviours to enhance immersion by imparting background characters with the necessary intelligence to form groups and navigate through highly populated virtual spaces. However, implementing an effective serious game or instructional environment within a virtual world requires the creation of characters that are not only visually perceived as behaving realistically on a large scale, but that are also capable of interacting with the user in a coherent and consistent manner, to maintain flow and immersion and thus effective learning transfer. For evidence of this relationship, see the works of Csikszentmihalyi and colleagues [6].

To enable and sustain this level of immersion and interactivity, we need to firstly define a framework capable of supporting such interactions and subsequently design a system which effectively implements these notions in a computationally and resource-efficient form. Our proposed approach integrates crowd modelling, intelligent characters and communicating agents into a composable and generalisable approach with potential application in a wide range of virtual environments and training simulations.

3 Levels of Interaction

In order to integrate the interactions between the player and the NPCs, we propose a novel framework called the Levels of Interaction (LoI). The LoI conceptually simplifies the interactions between the player and the NPCs. Graphically, the LoI can be represented as auras [11] of increasing complexity centred on the player’s avatar (figure 4) and based on a simple social space metric.

The first level aims to populate the area with an authentic crowd of characters. It provides a *living background* to facilitate the immersion of the player

in the environment. Characters evolving in closer surroundings of the player belong to the *interaction* level. These characters pay attention to the player and provide some basic interactions, comparatively to the next level of interaction. Finally, a character inside the *dialogue* level interacts with the player in a natural way, ultimately using speech recognition and synthesis.

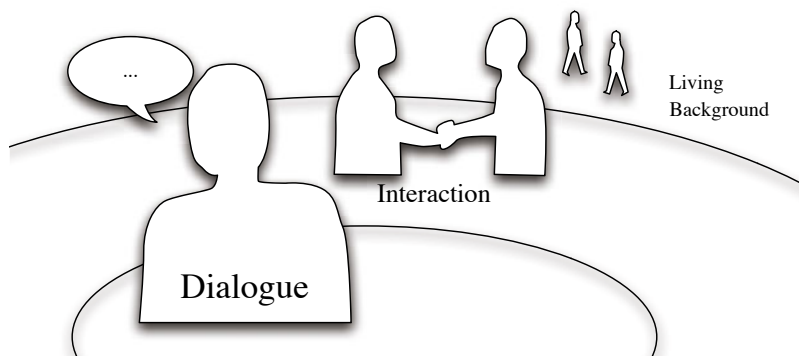


Fig. 4 The levels of interaction (LoI) technique allows different levels of complexity in the interactions between the player and the other characters, from an educational dialogue to a simple living background.

Characters are not assigned a specific level of interaction *a priori*. By default, they all belong to the background. But as the player moves and/or they move freely in the environment going about their business, they happen to get closer or away from the player and thus enter or exit the interaction or dialogue levels.

3.1 *Living Background*

The living background layer contributes predominantly to the believability of the environment. Kim and colleagues [16] utilise the term *believability* to define a two-layered model in an attempt to classify the various aspects of virtual environments of greatest concern when creating a real-world simulation. In this model, immersion plays a key role in creating a believable environment, in terms of the relationship between human sensory input and system output. The further human senses are exposed to external stimuli, such as background noise, Kim and colleagues argue, the lower the level of immersion and hence the less believable the environment. However, this is refuted to some extent by Robertson [32], who claims that “immersion should not be equated with the use of head-mounted displays: mental and emotional immersion does take place, independent of visual or perceptual immersion”, an opinion reinforced by Csikszentmihalyi and Kubey [6].

A simple out-of-place texture or inappropriate sound can have catastrophic effects for believability, as metrics of immersion such as the performance indications and cognitive surveys applied by Pausch and colleagues [26] demonstrate. With respect to the behaviour of avatars, this is similarly true – for example, if a character is perceived to be undertaking behaviour which is without context or motivation. The approach addresses these issues through not only the support of the living background level, but also by enabling the system to adapt to more sophisticated interaction requirements dynamically by elevating an interaction through the layers of the model. However, at the single crowd level, the believability challenge is addressed through two mechanisms. Firstly, by allowing the characters to adaptively evolve in an informed environment, they may respond and adapt to the behaviour of the player. Secondly, by endowing the characters with motivations and goals, the coherence of their behaviour may be assessed.

An informed or annotated environment contains semantic information, helping the characters to trigger the right behaviour regarding the context. Such information is useful for the general navigation of the crowd [9], defining routes, building entrances and associating areas with classes of characters to avoid incongruities – like a slave wandering in the *curia*, where the senators were meeting. Annotations may also be used for triggering specific behaviours related to specific places, like shops or street happenings. Although their behaviour is guided by these annotations, it is very important – for diversity’s sake but also for believability – that the characters maintain their own autonomy. Artificial Life (AL) provides techniques for enhancing their adaptivity. Indeed, increasingly ambitious approaches use AL-driven agents to simulate the varied and complex behaviours one can observe in a crowd, however relying on lightweight and simple systems. Such techniques, mostly inspired by Reynolds flocks [31], use physically-based systems [13] or cellular automata [17] to let properties emerge instead of scripting them manually. AL-crowds have revealed the ability to model coherent flowing and emergence of lanes, jamming behaviours at constricted points (e.g. doorways) and formation of groups. Finally, adaptive characters by definition have the ability to tailor their behaviour to a given situation, such that it is always coherent with regard to their context. This makes the characters opportunistic, as they can change their objective(s), for example going to the temple, depending on their mood or interest – attending a street exhibition or listening to an orator haranguing the crowd.

To maintain this coherency in time, characters also need to be motivationally-driven. As such, they are guided by internal motivations including internal needs, duties or a diary, so that everyone of them seems to have a reason to be part of the scene. If the player follows a character, he will witness a citizen living a coherent life. Furthermore, every character has a role and a psycho-sociologic profile [19]. The first defines strong relations between each other – a husband and wife, a master and slave(s), a centurion and soldiers, a merchant and customers. The latter defines soft relations, that

mostly depend on the situation – two friends encountering and waving at each other, or stopping their activity to chat, mendicants begging for money as a rich person passes by.

Although from the player’s point of view, the characters composing the background are merely able to walk or stop, getting closer will reveal in detail the interactions these characters are engaged in. This is explained in the next section, as part of the interaction level.

3.2 Interactions

Characters inside the interaction level are no longer part of the background since they provide a first degree of interaction with the player. This is revealed by a more complex behaviour providing a visual way to teach history, and the ability to dynamically respond to the player or engage them.

Inside the interaction level, the characters are rendered in more detail and animated with smoother motion captures in order to enhance their believability. Moreover, visually and historically more accurate motion captures can be used to depict scenes of craftsmanship, trading, art or rites, thus providing the learner with a first-hand perspective when learning history. Figure 5.a shows for example how virtual scenes of acting or dancing, taking place in the Roman theatre of Aspendos, might promote a better understanding of the contextualised social role such places occupied in the ancient societies [20]. The same technique could be applied to provide accurate depictions of the know-how of a baker, of the meticulous craft-work of a blacksmith or the social rules involved in a transaction between a merchant and a customer.

The interactive context of the game makes the use of a single lengthy motion-capture inappropriate to represent a complex behaviour. Rather, such a behaviour ought to be considered as an adaptive series of actions that the

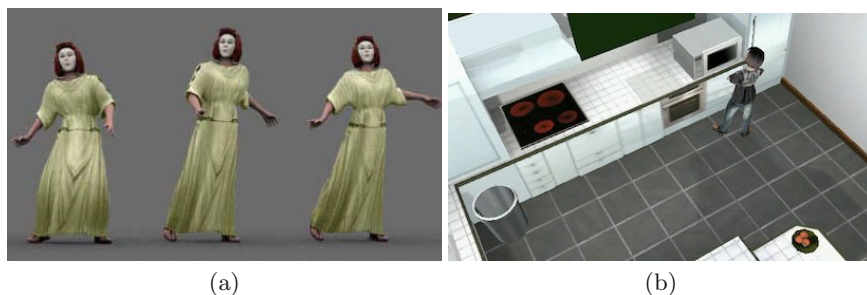


Fig. 5 (a) Accurate motion captures allow for learning transfer through interactive observation. (b) The VIBES architecture uses a classifier system to allow the automatic creation of adaptive complex behaviours from sequences of simple procedural actions. In this case, the character learns to cook a meal using basic ingredients and household appliances.

player would be able to interrupt dynamically. Agent technology provides a wide range of controllers for managing the complex behaviour of an animated character like a virtual human. ‘Hybrid’ controllers are well suited for this kind of task. They combine a set of basic behaviours to a database of knowledge representations within the same architecture, for the agent to be reactive enough to act naturally yet cognitive enough to act coherently. The ViBes architecture [33] (figure 5.b) has relevant features. It is firstly a hierarchical controller, so that every complex behaviour is built from the organisation of lower-complexity behaviours operating recursively until the lowest level of complexity, which consist in basic moves that can easily be rendered by motion captures. Additionally, the ViBes architecture uses AL techniques – namely, a classifier system – to ensure the adaptive sequencing of each behaviour, in such a way that any behaviour, no matter how complex, can be seamlessly interrupted by another, without leading to an unknown and hence non-displayable situation. As a result, a Roman character processing a series of actions towards a motivated goal can be interrupted by another character and start a new interaction scheme, at the end of which it would be able to return to their occupation.

The interaction level also provides the first degree of integration between the player and the crowd, by allowing proactive and non-verbal interactions with the other characters (figure 4). For instance, passing characters stare at the player, so that he does not feel like a ghost in the crowd. They could also greet the player, if for example the player has met them in the past. More importantly, these interactions provide a way to initiate a dialogue or give the user the incentive to initiate a dialogue. Indeed, non-verbal communication, such as waving or gaze behaviours, can play an important role in the initiation of a verbal interaction [27], for example by attracting the player’s attention. The player would then be free to step towards the character and thus enter the dialogue level.

3.3 *Dialogue*

No matter how many characters are included in the scene, the dialogue level, which involves a close-up interaction with a character, contains only one at a time. Since more details of this character can be potentially seen, this in turn suggests the requirement for a highly detailed representation, including for example, facial expressions, gestures or body stance [5].

When a character enters the dialogue level, a dialogue is triggered with the player. At this level, the role and capabilities of the character are well described by research into Embodied Conversational Agents [4], or ECAs. ECAs are 3D (figure 6.a) or cartoon-like (figure 6.b) characters that attempt to demonstrate many of the same capabilities as humans in face-to-face conversation, including the ability to produce and respond to verbal and non-verbal communication. ECAs may be endowed with a multimodal interface

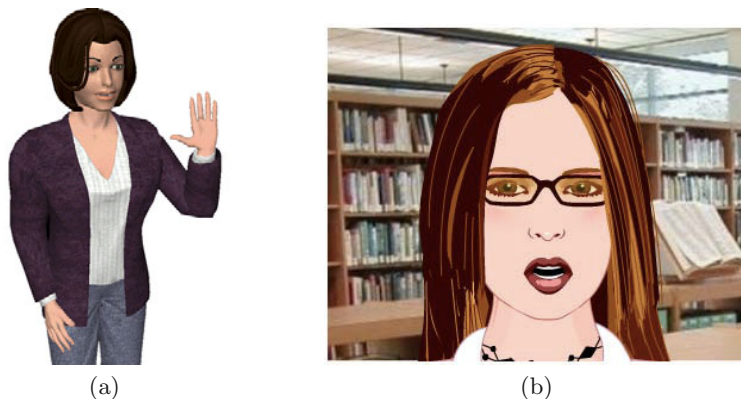


Fig. 6 Depiction of (a) Greta, an Embodied Conversational Agent [4], or ECA, and (b) Lillian the virtual librarian chatbot (©Daden Ltd).

that parallels their verbal output to other modalities, matching the emotional content of the dialogue or provide appropriate feedback. For example, the parameterised action representation [2] coincides specific expressions and gestures with speech, in order to bridge natural language and embodied action.

This level particularly supports situated, experiential learning. Therefore, the gaming aspect is largely supported by this level of interaction, and brought to the player via several quests, wherein each scenario is not explicitly defined, but rather built through the course of interactions with a variety of characters. By talking to different characters, the learner can interact with and reflect upon different aspects of the life in Rome and history and the differing roles and jobs of individuals within the society. As an example, the player may learn about military history while speaking to a praetorian or about diet when talking with a grocer.

These dialogic interactions are the very core of this history-teaching serious game, as known pedagogic strategies and techniques can be applied. Namely, the game could benefit from the experience of pedagogical ECAs in order to improve human learning performance. These specific ECAs are virtual instructors that autonomously train a human learner by applying appropriate pedagogical techniques during instruction [7] through scaffolded learning.

To summarise, the LoI framework proposes an elegant and convenient way to represent the types of interactions offered to the player through a single integrated system (figure 7). The next section presents a first implementation of the concept in the *Roma Nova* project, showing how technical issues have been overcome.

Level of Interaction	Living background	Interaction	Dialogue
Distance	Far	Mid-range	Near
LoD	Low	Medium	High
LoS	Low	Medium	High
Granularity	Crowd	Group	Individual
Agent properties	Reactive, Motivational, Opportunistic	Pro-active, Adaptive	Intelligent, Social, Conversational
Example capabilities	Navigation	Complex behaviour, Social attention	Facial expressions, Speech synthesis, Dialogue management
Computational resources	Low	High	Low

Fig. 7 This table summarises the different features of the characters, for each level of interaction. Note that in spite of higher realistic model and behaviour, the dialogue level is not the most costly in terms of computational resources.

4 Implementation

Although the levels of interaction are merely a conceptual architecture to frame the interactions between the player and the characters, their implementation relies on existing techniques, some of which are already implemented in game engines.

In order to prove both the relevance of the framework and its technical efficiency, a first implementation of the levels of interaction has been provided for the *Roma Nova* project. The implementation is based on the Unity game development tools [15]. Based on a recent internal study [29], Unity has shown to offer the best compromise between functionality and accessibility.

4.1 Computational Efficiency

A fundamental challenge addressed by the project is the effective modelling of interactions between the player and characters in a large environment, wherein a substantial number of entities are likely to interact in a non-linear and emergent way. Interestingly, this problem is very similar to modelling the visual and behavioural fidelity of the crowd. Many state-of-the-art techniques allow for the creation of visually realistic crowds. However, each technique is individually demanding in terms of computational resources and hence simplifications must be done in order to avoid disrupting graphical flow and fidelity and impeding perceived realism. Such techniques for ameliorating the impact of high-fidelity content on frame rate are known as *levels of detail* and *levels of simulation*.

The Level of Detail approach [18, 23] (LoD) involves creating, as part of a pre-process or at runtime, 3D representations of each object, capable of being resolved to varying levels of triangle resolution. There are currently four different frameworks for LoD: discrete, continuous, view dependent and hierarchical. Discrete LoD is a traditional approach that pre-processes several meshes of different complexity for each object. Continuous LoD adjusts the detail gradually. View-dependent LoD renders the portion of the object that is closer in higher detail. The Hierarchical LoD treats the scene as a single object and it simplifies the object according to the position.

The classical discrete implementation of the LoD has been preferred for *Roma Nova*. Although it requires more work from the modeller, we found at least one advantage using this technique as most game engines already include an implementation (like Unity for instance). The original mesh for the Roman character used in the game has been reworked using a modelling software in order to offer 3 levels of detail. A lower one, using a polygon simplification, was designed for distant crowd characters. Conversely, a higher mesh has also been designed for the dialogic interactions. Parts of the anatomy have been smoothed to avoid unpleasing angles whereas others have been re-created from scratch like eyes or lips, in order to fit the higher visual requirements of a close-up interaction. During the game, these meshes are switched seamlessly during the transitions between the levels of interaction, as explained further in this section.

The levels of simulation technique [30] (LoS) allows the computation time to be distributed according to the spectator's viewpoint, causing the quality to be enhanced in areas where attention is focused, e.g. around the viewpoint. The quality is decreased towards occluded or inconspicuous areas. For instance, the path-finding routine for some characters can be approximated

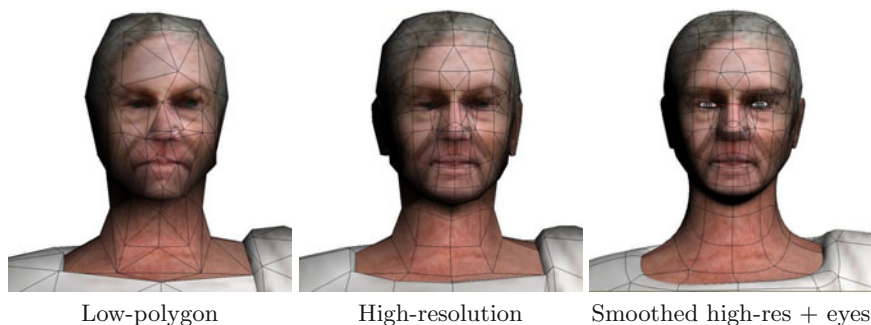


Fig. 8 Illustration of how levels of detail have been implemented for the *Roma Nova* project. Meshes of different complexity have been created on the basis of original the high polygon mesh. A polygon simplification has been applied to reduce the mesh complexity by 60% for the low-polygon version. The mesh has been smoothed and parts have been manually enhanced (eyes, lips, etc.) for the ultra high-resolution version.

and collision detection reduced in accuracy or removed. Whereas LoS are usually tricky to implement, as the behaviours have to be parameterised, the LoI framework provides a convenient simplification, allowing for one or more specific behavioural components to be attached to each level of interaction. Switching between two behaviours is managed exactly like switching the models, as explained below.

When modelling a crowd, the LoD and LoS techniques may be used in concert, such that NPCs closest to the camera are rendered and animated with the highest degree of fidelity (see figure 9). By comparison, NPCs in the far distance are typically rendered using lower resolution models and animations. Since the scene is intended as dynamic, LoD and LoS also address issues related to transitions between different levels. The key to implementing such an approach effectively is enabling seamless and imperceptible transitions between levels to the user and central to this is the seamless modification of the mesh [14] or the animation of the characters as it approaches or moves away from the observer.

Implementing the levels of interaction was quite straightforward as Unity allows to use boundary spheres associated with triggers (figure 10). The implementation relies on these spheres to trigger a change of the 3D model and/or the behavioural controller, therefore providing an implementation of LoD and LoS. Different controllers are loaded in real-time as characters enter or exit the levels of interaction. These controllers are then responsible for the degree of interaction allowed by the NPC.

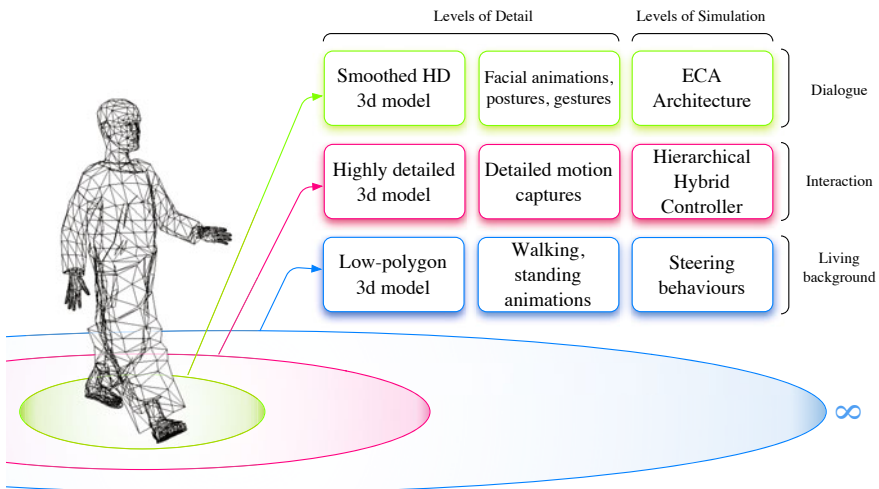


Fig. 9 Technically speaking, the LoI are implemented together with LoD and LoS, so that each level of interaction is related to a level of detail – one or several 3D models along with animations – and a level of simulation – one or several behavioural controllers.

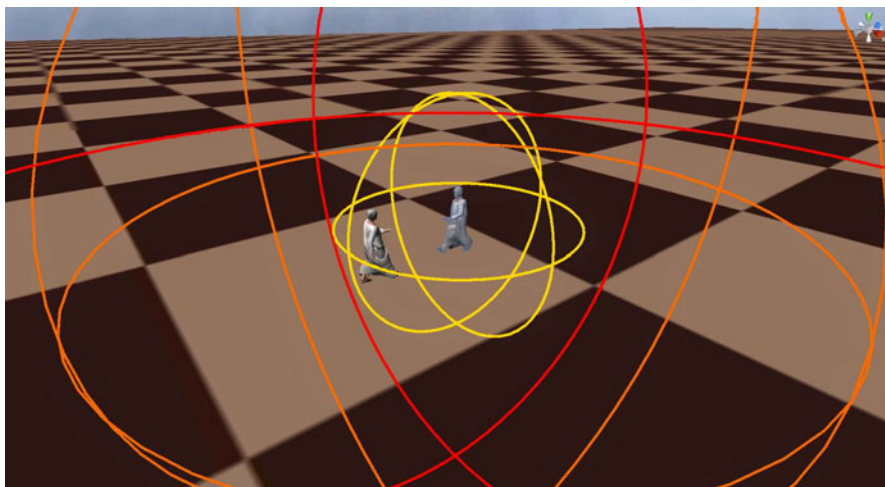


Fig. 10 Unity allows to use boundary spheres, on which triggers can be attached. The implementation of the LoI takes full advantage of this feature. Any NPC entering a sphere is triggered by the system and sent a message in order to update both its 3D model and behavioural controller. When leaving the sphere, a similar operation happen and the controller is dropped.

4.2 *Boundaries*

Whereas a critical aspect of the framework (as can be seen on figure 4) concerns the boundaries between each level of interaction, no indication of distance has been given so far. Although the boundaries have been considered as fixed radii centered on the player's avatar, the first implementation tends to suggest that it may only be a representation simplification and that a refinement of the definition of these concentric auras is required to accommodate a better understanding of human perceptual traits. Distance does not necessarily imply salience [8, 3] and the LoI approach may integrate with models for determining salience to provide a more perceptually-driven allocation of resources.

For instance, adjustable boundaries would provide dynamicism with respect to the graphical flow of the game, in addition to LoD and LoS techniques. Since the LoI integrates these techniques, as explained in section 4.1, they are directly responsible for the number of characters inhabiting each level of interaction. Assuming that a NPC showing a complex behaviour requires more computational resources than a background or a conversational character, as expressed in table 7, the LoI are responsible for the overall CPU and memory consumption of the game. To illustrate this, one can imagine for example that the number of characters allowed in the interaction level can be tweaked in real time by shrinking or expanding the corresponding

boundary in order to guarantee a constant framerate while maximising the player's immersion.

The dialogue level benefits from a specific rule. Although it is schematised as a radius, it is restricted to contain only one character at once, as current research on ECAs merely considers the verbal interaction with one agent only.

Another important aspect of the boundaries concerns the design of how they are meant to be crossed by the NPCs. Although switching the mesh and the controller is technically speaking an almost trivial operation, as explained in the previous section, the transition visually is likely to be perceived by the player and therefore act as an impediment to the immersion. Oscillatory phenomena and hysteresis must be foreseen as well, for instance when a NPC walks in the same direction as the player for a while, in and out of the interaction level. As a result, special care must be taken to ensure the visual and behavioural continuity between crowd and interacting characters.

Again, for the dialogue level, things are different given the specific "one character" rule. Also, the transition to a dialogue should not be seamless and graphically invisible but on the opposite explicitly notified to the player. The next section elaborates on how the interactions have been managed in the first implementation of *Roma Nova*. This includes an elegant solution for the dialogue transition.

4.3 Interactions

For the first implementation of *Roma Nova* we decided to temporarily defer the interaction level, focusing on the two remaining ones, and place specific emphasis on the transition from a crowd behaviour to the dialogue phase. In the current version, a crowd of Roman characters is introduced in the Forum, a highly detailed set of buildings that belong to the *Rome Reborn* model. The NPCs are wandering in the 3D environment between predefined points of interest, while the player is able to move freely (figure 11).

When clicked on by the player, a NPC changes its current target to the position of the player, and hence starts walking towards the player. When the NPC is close enough to enter the dialogue level, a series of actions is triggered by the engine:

- An animation is triggered to change the camera from a wide angle to a close-up perspective.
- The smoothed highly detailed version of the Roman character mesh (figure 8) is loaded to replace the low polygon version, along with the corresponding animations.
- The steering controller attributed to every background character is dropped and replaced by a simple ECA engine developed on purpose to play the scenario (figure 12).



Fig. 11 A crowd of Romans walk through the Forum to reach randomly distributed points of interest. The player, through his avatar, is free to summon any of them to start a dialogue.



Fig. 12 At the moment, the dialogues in *Roma Nova* consist of a simple textual interaction between the player and any character of the game. Future work intends to provide the conversational Romans with a fully-fledge ECA engine and content-rich dialogues.

At the end of the dialogue, the reverse set of operations is performed. The simple 3D model and the steering behaviours are re-applied to the NPC. It carries on its navigation to its previous point of interest as the computer memory is freed from the ECA engine and the high definition 3D model.

The next milestone will intend to introduce a first attempt of interactive behaviour. The characters from the crowd, as they get closer and pass by the player and thus enter the interaction level, will be endowed with a more realistic 3D model and gaze animations in order to visible respond to the player and thus acknowledge his presence. Characters already met by the player will be able to wave at him, hopefully encouraging the player to start a new conversation or to resume the previous one.

5 Conclusion

As well as a challenging opportunity to study how history could be taught *in situ* by means of a serious game, the *Roma Nova* project is an opportunity to explore the technical issues that must be solved to allow such a scope of interactions with a large number of believable virtual characters in an open and immersive environment. In this paper, our intention was to show that although this kind of environment has never been seen before in a simulation or a commercial game, it is less a question of computational efficiency than the lack of an adequate framework. Not only do the Levels of Interaction apply a divide-and-conquer strategy on the broad range of interactions, ultimately relying on current techniques to enable the implementation, but they also tailor the graphical, animation and behavioural levels to the interaction, so that the immersion of the player is always ensured.

Future work consists of the continuation of development work in order to explore and further refine the framework in the context of the *Roma Nova* case study. In parallel, further research intends to extend the framework to multiplayer games, where several players are immersed in the same environment, along with a crowd of NPCs and a human tutor.

Acknowledgements. The authors wish to express their thanks to the *Roma Nova* development team. All credits regarding the implementation of the project and the illustrations provided in the paper go to Kévin Sauvourel, Alexandre Spyropoulos and Guillaume Patinaud.

References

1. Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., Freitas, S.d.: Serious games in cultural heritage. In: 10th VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2009), VAST-STAR, Short and Project Proceedings, Malta, pp. 29–48 (2009)

2. Badler, N., Bindiganavale, R., Allbeck, J., Schuler, W., Zhao, L., Palmer, M.: Parameterized action representation for virtual human agents. In: *Embodied Conversational Agents*, pp. 256–284 (2000)
3. Beeharee, A.K., West, A.J., Hubbard, R.: Visual attention based information culling for distributed virtual environments. In: *VRST 2003: Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 213–222. ACM, New York (2003)
4. Bevacqua, E., Mancini, M., Niewiadomski, R., Pelachaud, C.: An expressive eca showing complex emotions. In: *AISB 2007 Annual convention, workshop Language, Speech and Gesture for Expressive Characters*, Newcastle, UK, pp. 208–216 (2007)
5. Bevacqua, E., Mancini, M., Pelachaud, C.: A listening agent exhibiting variable behaviour. In: Prendinger, H., Lester, J.C., Ishizuka, M. (eds.) *IVA 2008. LNCS (LNAI)*, vol. 5208, pp. 262–269. Springer, Heidelberg (2008)
6. Csikszentmihalyi, M., Kubey, R.: Television and the rest of life: A systematic comparison of subjective experience. *Public Opinion Quarterly* 45(3), 317–328 (1981)
7. Doswell, J.T.: Pedagogical embodied conversational agent. In: *IEEE International Conference on Advanced Learning Technologies*, pp. 774–776 (2004)
8. Dunwell, I., Whelan, J.C.: Spotlight interest management for distributed virtual environments. In: *Proceedings of 14th Eurographics Symposium on Virtual Environments*, Eindhoven, NL (2008)
9. Farenc, N., Boulic, R., Thalmann, D.: An informed environment dedicated to the simulation of virtual humans in urban context. In: *Eurographics*, pp. 309–318 (1999)
10. de Freitas, S., Neumann, T.: The use of ‘exploratory learning’ for supporting immersive learning in virtual environments. *Computers and Education* 52(2), 343–352 (2009)
11. Greenhalgh, C., Benford, S.: MASSIVE: a collaborative virtual environment for teleconferencing. *ACM Trans. Comput.-Hum. Interact.* 2(3), 239–261 (1995)
12. Guidi, G., Frischer, B.: Virtualizing ancient rome: 3d acquisition and modeling of a large plaster-of-paris model of imperial rome. In: *Videometrics VIII*, San Jose, California, USA, pp. 119–133 (2005)
13. Hegeas, L., Luciani, A., Thollot, J., Castagné, N.: A physically-based particle model of emergent crowd behaviours. In: *Proceedings of Graphicon* (2003)
14. Hoppe, H.: Progressive meshes. In: *SIGGRAPH 1996: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 99–108. ACM, New York (1996)
15. <http://unity3d.com/>
16. Kim, H., Di Giacomo, T., Egges, A., Lyard, E., Garchery, S., Magnenat-Thalmann, N.: Believable virtual environment: Sensory and perceptual believability. *Believability in Virtual Environments and Emotion Simulation* (2000)
17. Lightfoot, T., Milne, G.: Modelling emergent crowd behaviour. In: *The Australian Conference on Artificial Life (ACAL)*, pp. 159–169. University of New South Wales, Canberra (2003)
18. Luebke, D., Watson, B., Cohen, J.D., Reddy, M., Varshney, A.: *Level of Detail for 3D Graphics*. Elsevier Science Inc., New York (2002)

19. MacNamee, B., Dobbyn, S., Cunningham, P., O'Sullivan, C.: Men behaving appropriately: Integrating the role passing technique into the aloha system. In: Proceedings of the AISB 2002, pp. 59–62 (2002)
20. Magnenat-Thalmann, N., Foni, A.E., Cadi-Yazli, N.: Real-time animation of ancient roman sites. In: GRAPHITE 2006: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia, pp. 19–30. ACM, New York (2006)
21. Mam, J., Haegler, S., Yersin, B., Müller, P., Thalmann, D., Van Gool, L.: Populating Ancient Pompeii with Crowds of Virtual Romans. In: Proceedings of the 8th International Symposium on Virtual Reality, Archeology and Cultural Heritage - VAST (2007)
22. McDonnell, R., Larkin, M., Dobbyn, S., Collins, S., O'Sullivan, C.: Clone attack! perception of crowd variety. In: SIGGRAPH 2008: ACM SIGGRAPH 2008 papers, pp. 1–8. ACM, New York (2008)
23. O'Sullivan, C., Cassell, J., Vilhjalmsson, H., Dingliana, J., Dobbyn, S., McNamee, B., Peters, C., Giang, T.: Levels of detail for crowds and groups. *Computer Graphics Forum* 21(4), 733–742 (2002)
24. Panzoli, D., Peters, C., Dunwell, I., Sanchez, S., Petridis, P., Protopsaltis, A., Scesa, V., de Freitas, S.: Levels of interaction: A user-guided experience in large-scale virtual environments. In: IEEE 2nd International Conference in Games and Virtual Worlds for Serious Applications (VS GAMES 2010), Braga, Portugal, pp. 87–90 (2010)
25. Panzoli, D., Qureshi, A., Dunwell, I., Petridis, P., de Freitas, S., Rebolledo-Mendez, G.: Levels of interaction (loi): a model for scaffolding learner engagement in an immersive environment. In: Tenth International Conference on Intelligent Tutoring Systems: Bridges to Learning, Pittsburgh, pp. 821–823 (2010)
26. Pausch, R., Proffitt, D., Williams, G.: Quantifying immersion in virtual reality. In: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, pp. 13–18. ACM Press/Addison-Wesley Publishing Co. (1997)
27. Peters, C.: Direction of attention perception for conversation initiation in virtual environments. *LNCS*, pp. 215–228 (2005)
28. Peters, C., Ennis, C.: Modeling groups of plausible virtual pedestrians. *IEEE Computer Graphics and Applications* 29(4), 54–63 (2009)
29. Petridis, P., Dunwell, I., de Freitas, S., Panzoli, D.: An engine selection framework for high fidelity serious games. In: The 2nd International Conference on Games and Virtual Worlds for-Serious-Applications (VSgames 2010), Braga, Portugal (2010)
30. Pettré, J., Ciechomski, P.d.H., Maïm, J., Yersin, B., Laumond, J.P., Thalmann, D.: Real-time navigating crowds: scalable simulation and rendering: Research articles. *Comput. Animat. Virtual Worlds* 17(3-4), 445–455 (2006)
31. Reynolds, C.W.: Flock, herds and schools: a distributed behavioural model. In: SIGGRAPH 1987. *Computer Graphics*, vol. 21(4), pp. 25–34. ACM Press, Anaheim (1987)
32. Robertson, G., Czerwinski, M., Van Dantzich, M.: Immersion in desktop virtual reality. In: Proceedings of the 10th annual ACM symposium on User interface software and technology, pp. 11–19. ACM Press, New York (1997)

33. Sanchez, S., Luga, H., Duthen, Y., Balet, O.: Bringing autonomy to virtual characters. In: Ramos, F.F., Unger, H., Larios, V. (eds.) ISSADS 2004. LNCS, vol. 3061, Springer, Heidelberg (2004)
34. Shao, W., Terzopoulos, D.: Autonomous pedestrians. *Graph. Models* 69(5-6), 246–274 (2007)
35. Waldau, N., Gattermann, P., Knoflacher, H., Schreckenberg, M. (eds.): *Pedestrian and Evacuation Dynamics*. Springer, Heidelberg (2007)
36. <http://www.romereborn.virginia.edu>