

Human-in-the-Loop Simulation of a Virtual Classroom

Jesper Nilsson and Franziska Klügl^(✉)

School of Science and Technology, Örebro University, Örebro, Sweden
jrn1989.jn@gmail.com, franziska.klugl@oru.se

Abstract. As technology for virtual reality becomes more and more accessible, virtual reality based training becomes a hot topic as an application environment for multiagent systems. In this contribution, we present a system that connects a game engine with a BDI platform for simulating a group of listeners that may be believable enough for serving as a virtual audience for training gestures and body language while teaching.

1 Introduction

Virtual reality forms an interesting substrate for e-learning systems not just in distance learning applications or educational role-playing games. While military training systems were a kind-of early applications of risk-free and fully controllable environments, e-learning has discovered virtual environments enabling training experiences in a huge variety of domains (see [11] for a short review). In this contribution, we present a prototypical system for testing and training body language when speaking in front of an audience. The main addressees are teaching and research students who often have no chance to properly train their non-verbal behavior and style of presenting, experiencing how a group of listeners behaves differently in reaction to a particular posture or gestures. Rehearsing a talk in front of a simulated audience is without risk in the sense that the opinion or virtual agent is not important for the speaker. Also potential fatigue of the audience is not an issue. The group's reaction is fully controlled: if not wanted, there are no emergent group dynamics.

However, the audience behaviour cannot be trivial imposing challenges on the used agent architecture. When decoupling the simulation engine from the visualization game engine, we can integrate complex behaviour models that may generate required non-trivial agent behaviour with sophisticated visualization. Hereby, BDI agent architectures turned out to be a good starting point, offering a rather light-weight way of formulating flexible, and complex agent behaviour.

In the following, we will first discuss related work in more detail, also including a short analysis of requirements. After a description of the overall system, we indicate the particular way of using the BDI agent architecture integrating personality and abstract emotions in Sect. 5. Section 6 describes the set-up and results from a few initial experiments. The contribution ends with a short discussion of initial experiences and outlook to future work.

2 Virtual Reality-Based Training

Virtual Reality based training can be found in many variants, with commercial, academic and other applications. It is far beyond the scope of this contribution to attempt any form of a comprehensive review. Virtual reality may support training of special abilities or manoeuvres - such as in emergency medicine or for rescue operations. Even more than physical activities, Virtual Reality based training systems address behaviour. That means, a human immersed in the Virtual Reality system may train for example appropriate reactions in critical situations, e.g. when reacting to verbal attacks, managing a crowd or rescue management etc. In this contribution, we discuss a first prototype coupling agent-based simulation and a 3D game engine for training non-verbal aspects of lectures.

Also for this application area, one can find recent works. One example is *TeachLive*¹. It consists of a 3D virtual environment with “simulated” students which are actually controlled by human players invisible to the human interacting with them. Related studies are by Poeschl and Doering [22] and Slater et al. [25]. They deal with practising in front of virtual agents to overcome fear of public speaking. Poeschl and Doering discuss designing a realistic virtual audience based on observations of the behaviour of a real life audience. Slater et al. concentrate on the “presence response” as a metric for the efficiency of a virtual environment. Ideally the perceived level of presence should be equal to that of the real world. One interesting finding was that even with a low level of representational and behavioral fidelity in the virtual agents, the presence response was quite high. The training scenario in the TARDIS project [3] was a job interview. A human interacts with a virtual recruiter. The focus was not on an elaborated model of the recruiter agent, but on automated analysis of the test person’s state so that a maximum of useful advice could be given. Tools for emotion detection and other signals, developed in this project, can be very helpful also in our scenario for enabling the virtual students to react to multi-modal signals from the human “teacher”.

We did not find research that focused on modelling the simulated audience. Focus was put on how to use such a system, not how such a system can be built. In the following we concentrate on architectures to be used for human listener models as well as how to couple such an architecture to a visualization platform. We start by analysing requirements for the overall agent behaviour.

3 Human Behaviour Modelling

Considering artificial audiences, the following qualities can be postulated for believability partially derived from B. Hayes-Roth’s [12] and others’ discussion:

- There must be an intrinsic motivation for activities beyond reactions to external triggers. Believability of behaviour needs autonomy. A consequence is that agents need to possess pro-activity for behaviour without external trigger [5].

¹ <http://teachlive.org>.

- The virtual students exhibit heterogeneity. Behaviour should be typical and realistic for the situation. Consistent individual activities with different reactions by different agents to the same signal are necessary. An observer may attribute these differences to personality.
- Display of emotions forms an important feature for believability [7]. Heterogeneity not just concerns underlying personality types, but also relates to dynamism. Agents react in different ways to the same input over time. They shall possess some state that allows a situation to escalate. The overall dynamics shall emerge without being scripted. Emotions form a good basis for modelling this time-dependency and modifications of behaviour.
- The behaviour should not be fully predictable to a human observer. That means fully rational optimizing behaviour is not appropriate, but some reduced form of unpredictability might end up in interesting effects. So, simple variations related to for example how often a gesture is repeated or how fast the character moves, are not enough.
- The virtual characters shall also interact with each other and not just to the teacher. That means simply duplicating one agent is not enough.
- The behaviour just needs to appear believable for the experiments. The agents do not need to fully reproduce human behaviour and reasoning in other scenarios.
- The agents should be able to adapt their behaviour not just in reaction to what other agents did, but also with respect to the human “in the loop”. If there is a human presenter or teacher, the virtual agents should react to his/her gestures, actions, etc. It is not sufficient to pre-define more or less sophisticated behaviour, but the agent should display some feedback to the involved human.

It may appear surprising that aspects related to graphics and life-like visualization are not listed. We assume that consistent believable behaviour forms the essential ingredient – given a sufficient level of realism in visualization. This is also supported by our experiments. None of the human subject remarked existing flaws in the visualization, although there were obvious problems with e.g. agents standing in tables; the actual behaviour in interaction with the human was decisive.

We assume that a simple rule- or script-based architecture of the individual agents is not sufficient to produce behaviour satisfying those characteristics with reasonable modelling effort. Inventing a new architecture was out of question considering the huge variety of existing ones. In the following, we discuss agent architectures, that have already been used for controlling virtual characters in computer games or Virtual-Reality based training.

3.1 Cognitive Agent Architectures

Cognitive Architectures are special in the overall landscape of agent architectures that has been traditionally coined more using terms like deliberative, reactive, social or hybrid (see [32]). Cognitive architectures as used in general artificial

intelligence are not only used to build intelligent agents, but concretize theories of human cognition: “a cognitive architecture provides a concrete framework for more detailed modelling of cognitive phenomena, through specifying essential structures, divisions of modules, relations between modules and a variety of other essential aspects” [28]. That means they are based on assumptions and models how humans reason, organize memory etc. that are at least partially validated by experiments, interviews, etc.

Over the years, a number of those cognitive architectures have been suggested. In Soar cognitive processes as well as processes that determine behaviour are based on search in specific “problem spaces” that organize long-term memory in sets of rules. An elaborated decision making process either directly identifies rules relevant for the current context or selects another problem space to identify rule ranking, effect of actions, etc. This is combined with a learning process compiling new rules from solved problems.

Soar has been used for controlling virtual characters as opponents or collaborators for human players in virtual reality game and training settings for some time: QuakeBot [15] for controlling a non-player character in games and in later versions also for modelling opponents in military training scenarios [33] or TacAirSoar agents [29] for steering a plane in virtual manoeuvres with mixed human and agent formations.

To our knowledge, ACT-R [2] has not been used for virtual agents to the same extent as SOAR. ACT-R also forms a unified architecture aiming at human cognitive processes for reproducing phenomena known from cognitive psychology. Its overall set-up is modular reflecting hypotheses on modular brain functionality. It is based on hybrid approaches combining declarative and procedural knowledge processing, symbolic and sub-symbolic representations.

There are a number of other architectures for resembling human behaviour that are at least partially grounded in psychological literature. PSI developed by D. Dörner [10], as well as its implementation in MicroPSI [6], is based on sub-symbolic motive representation. OpenPSI [9] forms an implementation of the PSI theory based on OCP architecture combining uncertain logic representation and processing with approaches from computational linguistics, evolutionary learning and connectionist attention. CLARION [27] explicitly distinguishes between procedural knowledge that is represented sub-symbolically and declarative knowledge that is based on symbolic representation and explicit reasoning. A special focus is put on learning techniques to acquire these different knowledge types.

These cognitive architectures have in common that they combine various representations in different memory modules with complex reasoning processes to a unified architecture. However, for our aim it would have been not feasible to use one of those architectures given only limited resources. In the following, we will discuss architectures that are capable of producing believable behaviour of virtual characters without the claim to resemble real human cognition and reasoning.

Recently, more and more models involving social behaviour for capturing group dynamics have been published in agent-based simulation – an area in

which there is a clear need for transparent, theoretically grounded but feasible agent architectures. Modelling and simulation frameworks as presented in [31] specifically aim at testing crowd behaviour models with emergent group-level behaviour patterns. It might be an interesting line of research to test the applicability of such frameworks also in the domain addressed here.

3.2 Elements and Architectures for Believable Agents

In the area of Intelligent Virtual Agents, many approaches exist for creating life-like characters that generate believable behaviour [24]. Clearly, a focus is on appearance and visualization-related aspects, generating realistic facial expressions or gestures coherent with emotions that the agents shall express. Already [7], Bates et al. illustrated how many and which components an architecture for a generally believable agent needs, even if it is “just” a virtual cat. Emotional dynamics are hereby based on a sort of reaction to success or not of its action (measured along some “standards”). Attitudes towards other agents and events as well planned behaviour are also captured. Similarly also *PMFserv* was prominently developed for behaviour generation of believable virtual characters combining diverse functionality. Due to the wealth of incorporated components, this architecture appears to be as sophisticated as the cognitive ones described above.

For emotions and personality rather established and psychologically grounded models exist: Already the virtual cat of J. Bates et al. was based on the OCC model ([20] after [26]). The main idea is that emotions result from mental reactions to consequences of events, actions of agents (others and own) and aspects of objects. The OCC model elaborates positive and negative emotions depending on whether the consequence, etc. was expected or not and whether the agent sees it positive or negative. The OCC model is meanwhile widely used for creating emotions in virtual characters.

Personality appears to be a good approach to systematically express heterogeneity between virtual agents. Instead of randomly chosen parameter settings or behavior details, it better to combine them into a consistent “personality”. Due to its economic and coaching value, many basic concepts expressing traits that combine to a personality have been proposed. In virtual agents the model of the big five factors [18] appears to be prominent. This model assumes that five basic traits (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism) make up the personality of a human. It is purely descriptive based on a large volume of empirical work. Yet, there are only few agent architectures that integrates personalities in generic way.

BDI architectures seem to be agent architectures that are also used for virtual characters. The central idea is that goals and the plans for achieving those goals are two separate categories in the reasoning of the agent. Most BDI architectures – mostly descendants of PRS [13] – do not plan from first principles, but select and configure pre-defined activity descriptions that may contain sub-goals, actions, branching or loops explicitly capturing the procedural knowledge of the agents. There are many research works that use BDI architectures for

controlling virtual agents, such as [19,30], also several extensions for emotions or personalities have been suggested (e.g. [21]).

L. Antunes et al. [4] recently published very interesting conceptual considerations on enhancing goal-driven behaviour as used in BDI. They suggested to use more elaborated desire acquisition processes for generating socially realistic behaviour of simulated agents. An operationalisation of their approach – when available – might form an interesting alternative to the BDI-based architecture we used in the following. Clearly, in agent-based social simulation, we will find a wider range of agent architectures that might be suitable in our project context. A thorough search and evaluation will be part of our future work.

For our first tests, we selected to start with a basic BDI architecture as it allows for pro-active behaviour based on intrinsic motivations. Heterogeneity can be easily expressed. The architecture is flexible enough to also embed variations in the behaviour of an individual agent, as well as simple interactions between agents. Another important reason for selecting an BDI-based approach was the availability of convenient tools for implementing. However, although we could produce reasonable agent behaviour within a rather short time, the level of its genericness appears to be too high and we will continue testing other architectures. A relevant element of future research will hereby focus on architecture that involve predictions of how the human might react to actions.

4 Classroom Scenario

4.1 Information Flow

Figure 1 illustrates the overall set-up and information flow. Agents behaviour is visualized in some virtual reality display. Such a display can be a sufficiently large screen, a Cave or some modern VR device like Oculus Rift or similar. A human is immersed in the set-up and his or her behaviour is detected using suitable sensors. The sensor abstract raw data to information that steers an avatar representing the human in an agent-based simulation. The avatar is interacting with other simulated agents during a running simulation which is coupled to real-time for producing an appropriate impression for the advance of time. Each of the simulated agents connects to a virtual character in visualization. With today's technology, one may assume that all involved systems are sufficiently fast including reaction time of the simulated agents. Latency for communication between different systems is not relevant as usually reaction times do not need to be as fast as possible, but just result in a realistic impression.

4.2 Architectural Concept

There are two general approaches to connect an agent control architecture to a game engine, in case one does not want to re-implement the behaviour control within the game engine loop. The main difference consists in the representation of the environment: either there is an explicitly simulated environment that is

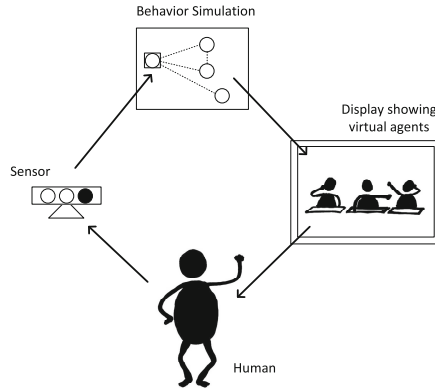


Fig. 1. Overall information flow in a system for training non-verbal behaviour with simulated listener agents

mapped to the virtual one including all agents that it contains being mapped to corresponding agents or there is just one representation of the environment in which complex agents with “external” reasoning capabilities are placed.

An elaborated middleware such as proposed by [16] or [30] transforms sensor data into percepts as input for the agent reasoning. The reasoning outputs high-level actions that are then transformed into control instructions for execution by the virtual characters’ body situated in the virtual environment. Sensing, action and interactions happen in the Virtual World that is accessed by elaborated interfaces between intelligent behaviour generation and the virtual characters’ body. Most systems mentioned in Sect. 3 use such a set-up. Those approaches are quite natural considering the analogy between physical and virtual reality environments and the physicalness of embodied agents.

We take a different approach that simplifies modelling interactions between agents and between agents and their environment keeping both on the same level of abstraction. Figure 2 illustrates the differences between both approaches. We basically double the environment and relevant entities for avoiding complex transformations in the interface between agents and their environment. The simulated agents interact with and in a simulated environment. All relevant events are triggered in the simulated environment. The virtual reality system serves merely for visualization purposes based on commands that inform the virtual character about what its corresponding agent is actually doing and translating that into gestures, facial expressions, etc. Clearly, those commands must be augmented with specific parameters for determining speed of a gesture, gazing directions, etc. Assuming that visualization can be created basically from existing 3d Object Models with given animations, handling behaviour within one system leads to simplified overall system development.

The commands for visual behaviour (movements, gestures, etc.) must contain specific symbols such as `WaveRightHand` or `ExpressJoy` triggering a particular

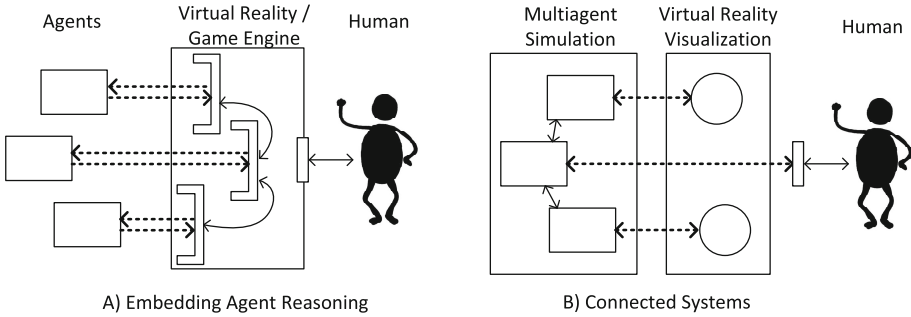


Fig. 2. Types of architectural for integrating multi-agent systems and virtual reality. The critical questionis where interaction between agents (as well as between agents and their environment) is actually happening.

gesture in a humanoid character model. The symbols must be understandable in a way that the visualization system can match it with a particular animation of the character model for e.g. raising the right arm and waving with it or setting up a facial expression for joy. Luckily, virtual characters and their animations are often reusable in the same domain. For example, the object models that we used in this work were developed at the University of Augsburg for a library of object models to be used in multiple projects.

In the overall system, the loop between visualization and simulation is not directly closed. Using a generic connection to the simulator for parsing strings sent to individual agents from software outside the simulator, we directly integrated sensor information into the simulation. Information form the sensor controls an avatar in the simulation, thus the simulated agents’ perception of and interaction with the teacher happens completely within the simulated environment. The avatar does not decide about actions that then manipulate the environment itself, it simply displays what the user did.

Such a general set-up was previously suggested for advanced visualization with the vision to create a platform for “immersive” validation [17]. In this project, a rich environment was not as essential as flexible and intelligent agent programs. We selected JASON [8] which appears to be currently the most elaborated, documented and easy to use open source BDI system.

4.3 Tools and Technicalities: JASON

JASON² [8] is a JAVA-based multiagent system platform around an interpreter of AgentSpeak(L) programs. AgentSpeak(L) is a BDI architecture with clear logics-based semantics. The interpreter takes care of handling agent’s belief and hierarchical plans and based on this, of action selection. A plan forms a piece of procedural knowledge consisting of specification of the triggering events (newly added/deleted percept, incoming messages or establishment or abandonment of

² <http://jason.sourceforge.net>.


```

/* Initial beliefs and rules */
posture(sit).
contentment(25).
mood(best) :-
    contentment(C) &
    C > 75.
mood(better) :-
    contentment(C) &
    C <= 75 &
    C > 50.
...
/* Initial goals */
/* Plans */
...
+spokenTo[source(Source)] : true
<- .drop_all_intentions;
!modifyContentment(25);
.wait(500);
lookAt(Source);
.wait(2000);
expressSadness(1);
lookAtCamera.

+hitByObject : true
<- .drop_all_intentions;
!modifyContentment(25);
.wait(1800);
lookAtPaper;
expressSurprise(1);
scratchHead(1);
.wait(2000);
lookAtCamera.
...

```

Fig. 3. Excerpt from a behaviour program of a nervous student agent

goals), the context in which the plan is applicable and the actual sequence of goals and actions that needs to be executed for achieving the plan. Goals can be the achievement a particular state or testing whether a particular condition holds. There is a number of built-in actions, for example for updating the belief base or sending messages to other agents. Domain specific actions can be added on the JAVA level elaborating the environmental model. Thus, an agent program consists of initial beliefs and goals, belief update rules and a set of plans that may be organized in multiple plan hierarchies. Figure 3 shows a snippet from an example behaviour of one of our agents. Thus, JASON provided the possibility to do fast prototyping of sophisticated agent behaviour in a human-readable form. Technically it works as an Eclipse plugin which enables easy JAVA-related extensions.

In contrast to full agent-based simulation platforms, predefined structures for the environmental model and the embodiment of the agents in JASON are quite rudimentary. Thus, we derived a specific environment class that manages additional representations of agents' bodies containing the relevant "physical" information such as position, orientation, current facial expression, gesture and focus point. Domain specific actions are dispatched by the environment and update the agents' body model. With a given frequency, updated information about what animations to display, etc. is send to the agents' corresponding virtual agent in the game engine. The human teacher has a simple, corresponding

agent in JASON that regularly updates information about the gesture that the human performed last receiving it from the sensor. Depending on such an incoming belief, the teacher agent sends corresponding messages to the other audience agent.

4.4 Tools and Technicalities: Horde3D Game Engine and Connection Components

Visualization was done using the Horde3D Game Engine³ with extensions that we developed previously for connecting to an agent-based simulation platform as sketched in [14]. The Horde3d Game Engine was and is developed at the University of Augsburg, Germany. It is an open source, light weight and conceptually clean game engine on top of a graphics engine with the same name. The component responsible for connecting the external behaviour control to the game engine's characters is based on proprietary strings sent via a socket connection containing information on what needs to be changed in the characters' display. The Horde3D component parses the string and executes the changes. Control of delays and parameters of execution (speed, repetitions of gestures, etc.) are determined by the behaviour control. Also the component that recognizes human gestures via the Kinect sensor is an existing component of the Game Engine.

5 Listener Behaviour

The central aspect of our research was not just doing a hardware and software set-up but actually creating believable audience behaviour using this for testing how far one can come with such simple means.

5.1 Action Repertoire and Interactions

In the agent plans, actions could be used that correspond to gestures and other actions displayed by the 3d characters in the visualizing Game Engine. Agents could stand up, turn, write, nudge or approach their neighbours, shake their heads, show a number of gestures such as pointing to the watch, move the hand to the head in a thinking gesture, etc. Also facial expressions for emotions were included, yet not really well recognizable in the visualization. Other students in the classroom however, could not miss-understand agents' actions and facial expressions as they perceived within the (high-level) simulation platform. Thus, no interpretation was necessary; they were transmitted as clear symbols in messages between agents. Only the human teacher could miss-interpret what the agents displayed.

In a similar way, a set of pre-defined gestures could be recognized via the Kinect sensor. Once the gesture was sent to the agent simulation system, all agents had the same understanding. There were a number of miss-interpretations

³ <http://www.hcm-lab.de/projects/GameEngine/>.

of human gestures, as recognition was not perfect, but in those situations all audience members miss-understood what the human wanted to convey in the same way.

5.2 Emotions and Personalities

Two options were used to individualize the audience agents: emotions and personality. Inspired by the OCC model, we designed every agent to have a numeric variable expressing some general form of “contentment” with its individual situation. This can be basically seen as a one-dimensional appraisal model of emotion: Each event, that means each perception or incoming messages has an effect on the level of this variable, its value is increased or decreased. So for example, if an agent is approached by one of its neighbours, the contentment level is modified by certain number. The number, and actually whether it is increased or decreased depend on the agents’ particular personality. The contentment forms both a modulator of how often a gesture action is done or also determining which gesture is selected as a reaction to an event. For our implementation in Jason, this means: The numeric variable is abstracted into a categoric statement like “mood(bad)”, which is then used as context for plan selection.

Yet, integrating emotional behaviour alone does not produce sufficient individualism. Agents need to exhibit more particularities than what can be expressed by for example different start values for the contentment level. The interaction in this project is not rich enough for that compared to conversation-oriented scenarios as interaction is solely based on gestures and postures of the teacher, and on approaches or attacks from other agents. As discussed above, equipping the agents with personality appears to be the gold standard for individualism. We selected only a few personalities that we assumed to challenge a teaching situation. Without further elaboration, we choose to model agents on the positive and negative extreme of extraversion, neuroticism and non-agreeableness. Together they might form a “good” audience for training non-verbal reactions to student behaviour. Yet, this is clearly a decision that needs further grounding in pedagogical and psychological research. The personalities cause differences in the agent programs, formulated in a quite ad hoc way. Simulated students with different personalities react to different incoming events, for example a “timid” student reacts to hardly any gesture and also continues ignoring its neighbours approaching it. A nervous student will also not react to its neighbours in a meaningful way, but becomes more and more uneasy, the longer the teacher is inactive, eventually disturbing other agents. A hostile student agent shows aggression towards its neighbours and reacts with an angry expression to many interactions from other students, but also to most gestures of the human teacher. Extravert students were designed so that they interact frequently with their neighbours and are quite active especially when the human is inactive. We added more than one hostile and extravert personality making additional differentiation in parameters, durations but also in a few events that they react that others ignore. In the final scene there are one timid, one nervous,

two hostile and four extravert students. If there is no activity from the teacher identified, the situation escalates.

A critical aspect in developing those agents' personality and behaviour was the parametrisation of the dynamics of the "contentment" variable. There are clearly a lot of ways to extend and improve the agent programs for the different agent personalities. Also the repertoire of gestures that the student agent react to can be extended as well as the reactions could be much more sophisticated. We see that as a starting point for further developments in collaboration of the pedagogic department in Örebro. Despite of the shortcomings in the behaviour definitions, we tested how human subjects would react to the virtual agents audience and conducted experiments for getting an impression how humans would rate the realism of the virtual audience.

6 Experiments

6.1 Experimental Set-Up

Using the configuration described above, we exposed 16 subjects to two sub-experiments. In the first one, the task was to observe a more and more escalating situation without teacher interaction (the sensor was not turned on, so the teacher agent was observed to be inactive by the student agents over the duration of that experiment). So, the subject could concentrate in observing and evaluating the behaviour of the simulated audience. Figure 4 shows how the situation looked like at the start of the experiment and how it could look like at the end. The behaviour of the agents is not scripted, so the final situation was slightly different in each of the experiments even without interactions with the human teacher. For the second experiment, the subjects were informed about the gestures that the system could recognize, yet not how the agents would react to those gestures. The subjects' task was to keep the groups' attention. Both experiments lasted 5 min.

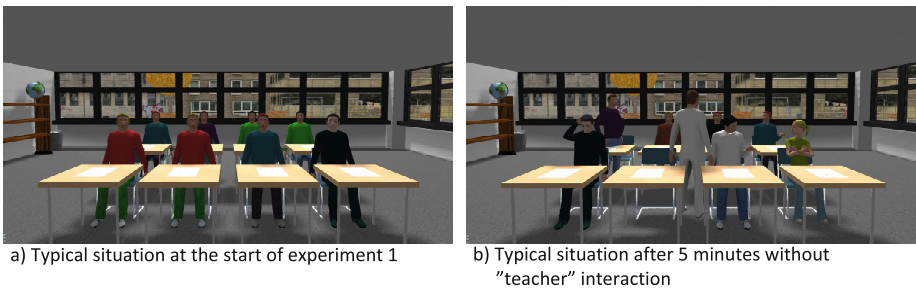


Fig. 4. Scenarios during the observation experiment. These are screenshots from two different runs; gender and cloth of the virtual agents were randomized.

After each sub-experiment the subjects were asked to fill a questionnaire. There were only slight variations between the two questionnaires: They contained

three groups of questions: (1) General question on the perceived realism of the scenario; (2) questions on the experienced emotions during participating the experiment and (3) open questions for feedback. The idea behind those questions was first to get a general evaluation and to check whether the subjects would perceive variations between the simulated students in a way that could allow them to identify the students' personalities. Emotional reactions of the subjects indicate that they experienced some form of presence. The rationale behind giving twice almost the same questionnaire was to find out whether interaction changes the evaluation that the subject gives.

The experiment were performed with 16 subjects recruited mostly from PhD students, Postdocs and Lecturers at the computer science, mathematics and technology departments at Örebro University. 6 of the subjects were female, 10 male. All had teaching experience; 10 subjects stated experience with games involving some form of motion capturing.

6.2 Results

In the following we only show the most interesting results concerning overall believability (Fig. 5). There is a tendency that subjects find the audience more realistic when they just observe rather than interact. We observed that some of the subjects spent some time in “testing” gestures systematically.

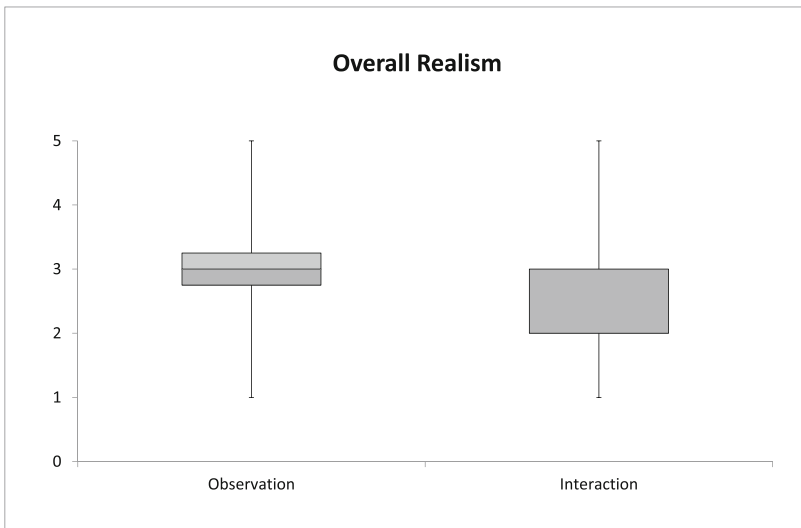


Fig. 5. Answers on the question “How would you rate the overall realism of the agents’ behaviour?” with 1 as “Not realistic at all” to 5 as “very realistic”.

Interestingly, some subjects also expressed emotional reaction in the first experiment in which they were told not to interact with the audience. So, some

limited form of presence could be seen because the students reacted to the missing actions of the teacher - more and more interactions among the audience occurred, most of the agents turned away from the teacher which was what the subjects expected them to do as the teacher remained inactive. Some teachers felt stressed as they were not allowed to intervene. Despite of the limited gesture repertoire, the most subjects felt that they could influence the students and were in control of the situation. In the open questions, many subjects expressed the need for more modes of interaction.

Heterogeneity of student behaviour was well recognized - both between students as well as over time; yet hardly any subject could identify the particular personalities displayed.

7 Discussion and Future Work

The experiments – although far from providing us with statistically significant results – show that it is possible to set up such a human-in-the-loop system with a virtual audience using rather simple, existing technology. Clearly, we just tested for believability of agent behaviour. For actually testing potential training effects, much more has to be invested from improvements in the gesture recognition to visualization that is fine-grained enough for enabling the reliable perception of the characters' emotions. The behaviour and especially the reaction to the teachers' actions need to be improved to be more realistic; Currently only limited tuning was done. We did actually not expect that subjects rate realism higher than they did, yet we assumed positive effects of enabling interaction. The question remains, what is the minimum necessary level of plausibility of audience behaviour and interaction so that the non-verbal behaviour training for teachers is useful? Addressing this question must be clearly the next step. A better selection of gestures and actions that the teacher can do, would be an essential for that.

Whether more realistic group behaviour (as proposed in [23]) or augmenting BDI reasoning with reasoning about emotions (e.g. [1]) have to be integrated, is a consequence of this next step. Is it necessary to develop simulated students who not just react to gestures, but intentionally drive the human teacher mad? Technology-oriented research in this project has set a first step, but without a wealth of tests and extensions grounded in pedagogical and psychological research, it is done in vain.

Acknowledgement. The authors would like to thank Elisabeth André and her group at the University of Augsburg - in particular Michael Wissner - for their support with the Horde 3D Game Engine as well as with providing the 3d Object models for the virtual students and the scenario visualization.

References

1. Adam, C., Lorini, E.: A BDI emotional reasoning engine for an artificial companion. In: Corchado, J.M., et al. (eds.) PAAMS 2014. CCIS, vol. 430, pp. 66–78. Springer, Heidelberg (2014)
2. Anderson, J.R.: Rules of the Mind. Erlbaum, Hillsdale (1993)
3. Anderson, K., et al.: The TARDIS framework: intelligent virtual agents for social coaching in job interviews. In: Reidsma, D., Katayose, H., Nijholt, A. (eds.) ACE 2013. LNCS, vol. 8253, pp. 476–491. Springer, Heidelberg (2013)
4. Antunes, L., Nunes, D., Coelho, H.: The geometry of desire. In: Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014, Paris, France, pp. 1169–1172, May 2014 (2014)
5. Avradinis, N., Panayiotopoulos, T., Anastassakis, G.: Behavior believability in virtual worlds: agents acting when they need to. SpringerPlus **2**, 246 (2013)
6. Bach, J.: MicroPsi 2: the next generation of the MicroPsi framework. In: Bach, J., Goertzel, B., Iklé, M. (eds.) AGI 2012. LNCS, vol. 7716, pp. 11–20. Springer, Heidelberg (2012)
7. Bates, J., Lyall, A.B., Reilly, W.S.: An architecture for action, emotion, social behavior. In: Castelfranchi, C., Werner, E. (eds.) MAAMAW 1992. LNCS, vol. 830, pp. 55–68. Springer, Heidelberg (1994)
8. Bordini, R.H., Hübner, J.F., Wooldridge, M.: Programming Multi-agent Systems in AgentSpeak using Jason: A Practical Introduction with Jason. Wiley, Chichester (2007)
9. Cai, Z., Goertzel, B., Zhou, C., Huang, D., Ke, S., Yu, G., Jiang, M.: OpenPsi: a novel computational affective model and its application in video games. Eng. Appl. Artif. Intell. **26**, 1–12 (2013)
10. Dörner, D.: Bauplan für eine Seele. Rowohlt, Reinbeck (1999)
11. Eschenbrenner, B., Nah, F.F., Siau, K.: 3-d virtual worlds in education: applications, benefits, issues, and opportunities. J. Database Manag. **19**(4), 91–110 (2008)
12. Hayes-Roth, B.: What makes characters seem life-like. In: Prendinger, H., Ishizuka, M. (eds.) Life-Like Characters, Cognitive Technologies, pp. 447–462. Springer, Heidelberg (2004)
13. Ingrand, F.F., Georgeff, M.P., Rao, A.S.: An architecture for real-time reasoning and system control. IEEE Expert Intell. Syst. Appl. **7**(6), 34–44 (1992)
14. Klügl, F., Wissner, M., Timpf, S., Andre, E.: Bridging virtual world visualization and multiagent simulation (extended abstract). In: Proceedings of the 11th Scandinavian Conference on Artificial Intelligence, Trondheim, May 2011, pp. 191–192 (2011)
15. Laird, J.E., Duchi, J.C.: Creating human-like synthetic characters with multiple skill levels: a case study using the soar quakebot. In: AAAI 2000 Fall Symposium Series: Simulating Human Agents, November 2000 (2000)
16. Lee, J., Baines, V., Padget, J.: Decoupling cognitive agents and virtual environments. In: Dignum, F., Brom, C., Hindriks, K., Beer, M., Richards, D. (eds.) CAVE 2012. LNCS, vol. 7764, pp. 17–36. Springer, Heidelberg (2013)
17. Louloudi, A., Klügl, F.: Immersive face validation: a new validation technique for agent-based simulation. In: Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 1255–1260 (2012)
18. McCrae, R.R., John, O.P.: An introduction to the five-factor model and its application. J. Personality **60**(2), 175–215 (1992)

19. Norling, E.: On evaluating agents for serious games. In: Dignum, F., Bradshaw, J., Silverman, B., van Doesburg, W. (eds.) *Agents for Games and Simulations*. LNCS, vol. 5920, pp. 155–169. Springer, Heidelberg (2009)
20. Ortony, A., Clore, G.L., Collins, A.: *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge (1988)
21. Parunak, H.V.D., Bisson, R., Brueckner, S., Matthews, R., Sauter, J.: A model of emotions for situated agents. In: *Proceedings of AAMAS, Hokodate, Japan*, pp. 993–996 (2006)
22. Poeschl, S., Doering, N.: Virtual training for fear of public speaking—design of an audience for immersive virtual environments. In: *Virtual Reality Short Papers and Posters (VRW)*, pp. 101–102. IEEE (2012)
23. Prada, L., Paiva, A.: Teaming up humans with autonomous synthetic characters. *Artif. Intell.* **173**, 80–103 (2009)
24. Prendinger, H., Ishizuka, M. (eds.): *Life-Like Characters*. Springer, Heidelberg (2004)
25. Slater, M., Pertaub, D.-P., Barker, C., Clark, D.: An experimental study on fear of public speaking using a virtual environment. *Cyberpsychol. Behav.* **9**(5), 627–633 (2006)
26. Steunebink, B.R., Dastani, M., Meyer, J.-J.C.: The OCC model revisited. In: Reichart, D. (ed.) *Proceedings of the 4th Workshop on Emotion and Computing, KI 2009, Paderborn* (2009)
27. Sun, R.: Learning, action and consciousness: a hybrid approach toward modelling consciousness. *Neural Netw.* **10**(7), 1317–1331 (1997)
28. Sun, R., Coward, L.A., Zenzen, M.J.: On levels of cognitive modeling. *Philos. Psychol.* **18**(5), 613–637 (2005)
29. Tambe, M., Johnson, W.L., Jones, R.M., Koss, F., Laird, J., Rosenbloom, P.S., Schwamb, K.: Intelligent agents for interactive simulation environments. *AI Mag.* **16**(1), 15–39 (1995)
30. van Oijen, J., Dignum, F.: Towards a design approach for integrating BDI agents in virtual environments. In: Vilhjálmsson, H.H., Kopp, S., Marsella, S., Thórisson, K.R. (eds.) *IVA 2011*. LNCS, vol. 6895, pp. 462–463. Springer, Heidelberg (2011)
31. Wijermans, N., Jorna, R., Jager, W., van Vliet, T., Adang, O.: CROSS: modelling crowd behaviour with social-cognitive agents. *J. Artif. Soc. Soc. Simul.* **16**(4), 1 (2013)
32. Wooldridge, M.: *An Introduction to Multiagent Systems*, 2nd edn. Wiley, Chichester (2009)
33. Wray, R.E., Laird, J.E., Nuxoll, A., Stokes, D., Kerfoot, A.: Synthetic adversaries for urban combat training. In: *Proceedings of the 16th Conference on Innovative Applications of Artificial Intelligence, IAAI 2004*, pp. 923–930. AAAI Press (2004)