Building Exposure: Synergy of Interaction and Narration through Social Channels

Bruno HERBELIN - Michal PONDER - Daniel THALMANN
Virtual Reality Laboratory – EPFL
Lausanne – Switzerland
{firstname.name}@epfl.ch

Abstract
The ability to build and control exposure through a seamless synergy of interaction and narration is strongly required by a new type of immersive VR training and therapy systems. This paper presents a practical approach for immersive VR training and therapy applications based on Interactive Narration Space (INS). The approach satisfies both interaction and narration requirements through the use of a high level Social Channel interaction paradigm. By introducing the Social Channel, we aim to minimize the contradictions between control over the story required by trainers/therapists, and interaction required by trainee/patient.

Keywords
Exposure, Virtual Reality, interaction, narration, social, virtual human.

1. INTRODUCTION

Growing availability of powerful graphics hardware, able to support high-end immersive VR simulations, brings fresh interest to a broadening spectrum of organizations in the new potential offered by modern VR-based training and therapy applications. Originally, VR training systems were focusing on skills training involving operation of strictly technological entities (airplanes, nuclear power stations, complex hardware devices, etc). Concerning VR therapy, applications were even more limited, if not marginal. Only recently did advanced virtual character simulation technologies reach a level of believability that extends the spectrum of VR applications to other domains such as medicine, emergency or psychotherapy, i.e. domains that require interaction with Virtual Environments featuring believable virtual humans.

With this powerful hardware and software technologies at hand, we face a problem: how to allow trainers and therapists to author and control the exposure required by these new application domains? An even bigger challenge is the migration from the well-known interactive scenes to the interactive narrations which trainees and patients should be exposed to.

In the following paragraphs we describe a methodology for combining narration and interaction in order to generate a required exposure with the means of an immersive VR simulation system.

2. VR IN TRAINING AND THERAPY

Overview

Immersive VR is able to provide rich, interactive and engaging training context that in reality would be too dangerous, too expensive or simply impossible to access. It provides a '*learning by doing*' medium that allows enhancement of the learning process by first-person experience [2]. An overview of concrete applications and techniques used to implement VR training systems in different domains is given in [8]. Successful use cases as well as new attempts can be found in a very broad range of domains such as space and aviation [7], military mission training [9], medicine (e.g. gunshot wound surgery [4], bone dissection surgery [6], virtual endoscopies[16]), emergency [19], health emergency [18], etc.

Many of the VR training systems realized can be categorized as skills training. In contrast, situation training systems aim at helping the trainee to develop certain psychological abilities required to face reality. While skills training systems focus on '*learning by doing*', situation training aims at '*learning by living*' through the virtual experience. Such a simulation imposes on the trainee the need of swift and precise assessment of the situation and rapid decision-making. These are usually combined with uncertainty of choice, adding in turn to stress, making it difficult to use theoretical knowledge gained through the regular in class course.

In psychiatry, Cognitive and Behavioral Therapy (CBT) uses exposure to fearful situation and anxiety-provoking stimuli to treat patients suffering from phobias (i.e. many anxiety disorders like social phobia, acrophobia, arachnophobia, etc.). According to Riquier et al.[17], “the goals are to decrease the avoidant behavior linked to anxiety disorders and to cope with provoked anxiety”. Virtual Reality entered the mental health field some years ago to determine if Virtual Reality Exposure (VRE) may be an alternative to standard in vivo exposure for wide phobias. Several studies reported that VRE is as effective as exposure in vivo on anxiety and behavioral avoidance therapy [15, 10]. Claustrophobic and acrophobic applications of VRE have been developed as a tool for mental health therapists [1]. VRE therapy has been successful in reducing fear of heights [14]. VRE has also been applied to the case of Social Anxiety Disorder: Pertaub et al. [13] showed that social anxiety can be induced by a virtual audience and that the degree of anxiety experienced is directly related to the type of virtual audience feedback received by the speaker.

Common Functional Requirements

At first sight, situation training and psychological therapy seem to represent two distant VR simulation domains. However, from the VR perspective one can observe some important correlations that allow for a common definition of the system functional requirements:

- *Specialist’s support*: the simulation is driven by a trainer or a therapist who supervises the simulation (authoring and control),
- *Subject’s performance*: the immersive experience needs to be interactive for the subject while at the same time the operator needs to have control over the evolution of the story in order to assure exposure of the trainee/patient to the key scenario elements,
- *Cognitive impact*: the subject must remember the experience as lived-through.

To sum up, the main objective is to guarantee that the trainee/patient is able to interact with the environment, while at the same time follow a well-defined narrative storyline of pedagogic/therapeutic purpose, featuring certain invariant elements required for the success of the training/therapy. Form the trainee/patient perspective, the system should be interactive, while from the trainer/therapist standpoint it should allow for pedagogical/therapeutic narration. The immersive VR system should resolve this contradiction by finding an optimal compromise.

3. PROBLEM STATEMENT

In this section we define the key concepts that will be used in the statement of the main requirements for an immersive VR system allowing for authoring and runtime control of exposure required by training and therapy applications.

3.1. Key Concepts

Exposure

Extending the strictly psychiatric concept, one can consider *exposure* as a generic situation where a person lives an experience controlled by the will of someone else. The *exposure* can be real or built artificially. Immersive technology aiming at generating a Sense of Presence among the simulation participants provides very powerful means that allow for design and control of *Virtual Reality Exposure*.

Interactive Narration Space

In contrast to the interactive space focusing on interaction with elements of the 3D scene, the Interactive Narration Space (INS) denotes a narrative environment allowing for interactive interventions from the participant’s side that affect the story evolution. INS shall provide proper semantics allowing for expression and execution of interactive stories.

Social Channel

In the context of VR simulation, the Social Channel (SC) denotes an interaction paradigm relying on a direct and human-like communication between the participant and Virtual Humans (VH’s). In other words, interaction of the participant with the VR environment is mediated through VH’s, who execute orders, ask questions, speak about simulation state, etc. Their presence is seamlessly combined into the fabric of the pedagogical/therapeutic story, hence they are not perceived explicitly as an “interaction method”. The SC defines negotiation based interaction, where the VH’s have the following roles:

- VH's can refuse to execute orders that would push the story into undesired directions (from the pedagogical/therapeutic point of view). However, such behaviors are not frustrating for the participant, since they take place in a social interaction context (as opposed to human-machine interaction, which assumes absolute superiority of human over machine).
- VH's encourage, suggest and prompt the user to take certain actions. This allows the trainers/therapist to guide the participant through the interactive scenario.
- Finally, in case of lack of interaction/cooperation from the participant's side, the VH's may make decisions by themselves. This again is acceptable from the social behavior point of view and may even have stimulating effects on the participant, encouraging him to take an active role instead of following someone else.

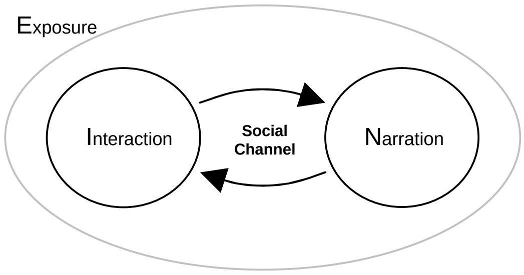


Figure 1. Social Channel (SC) combining interaction and narration

Negotiation and mediation is a natural interaction mode for the trainee/patient, while the trainer/therapist can control the narration into pedagogically/therapeutically meaningful directions : the SC forms the bridge that allows compromising and masking the inherent contradiction between interaction and narration (see Figure 1).

3.2. Requirements

The main objective is to provide trainers and therapists with an immersive VR system that would allow them to author and control exposure. On top of the VR immersive technology, we propose a practical implementation of the Interactive Narration Space concept in order to combine guided narration (as needed by trainers/therapists) and interaction with the virtual environment (as perceived by trainees/patients). In the context of INS, interaction paradigms need to be carefully elaborated in order to assure diversity of the interactive stories.

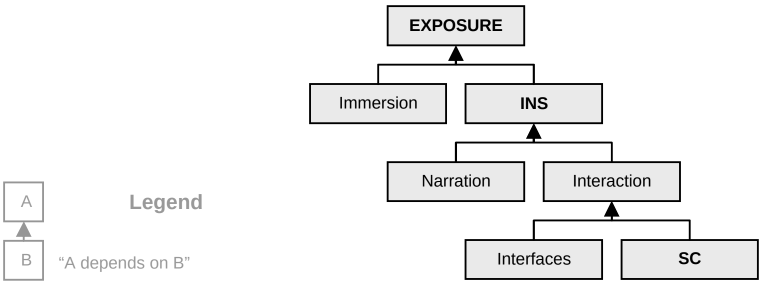


Figure 2. Building exposure: highlighting dependencies on the Social Channel side

It is well known that the use of interactive VR devices and interaction paradigms heavily constrains the spectrum of application domains. Thus it is assumed that the SC-based interaction paradigm will play the main role. SC may avoid common problems related to human-computer interaction, and in addition it is very compatible with the requirements of the INS. The layered dependencies between the functional elements are presented schematically in Figure 2.

4. APPROACH

4.1. INS Semantics

The basic idea of this approach is to map both narration and interaction onto a single decision tree. Using tree semantics, all possible narrations are expressed as different paths from the root (beginning) to the leaves of the tree (ends). On runtime, the execution evolves though the scenario steps which can be regarded as interaction points changing the story direction, but only within limits specified by the authors (trainers/therapists).

The semantics makes a distinction between *scenario step*, *decision group*, *decision* and *action*. From the bottom, an *action* is a predefined non-interactive simulation sequence (usually combining animation, sound, and change of state). An *action* is linked to a *decision*, and is triggered when the latter is made. A *decision group* is a set of *decisions* (choices) the user can make at a certain time (in a node of the decision tree) in any order. A *decision group* must contain at least one *decision* and needs to specify its *completion ratio*. The *completion ratio* provides number of *decisions* that need to be triggered within a given *decision group* in order to be able to proceed to the next *scenario step*. The *decision group* and *completion ratio* allow expressing situations where multiple actions need to be performed but where their order is not important. A *scenario step* is a set of *decision groups* defining different branches in the scenario tree. There must be at least one *decision group* in a step called *default decision group* (that will define the default path). A list of optional context-related *decisions* can also be provided to offer more possibilities, but these do not lead to another *scenario step*. On top of it, the *scenario* has to reference only the first scenario step (root of scenario tree) and an optional list of always-available *decisions* (e.g. “Look at my watch and read time”). Figure 3 gives a schematic example of a scenario structure.

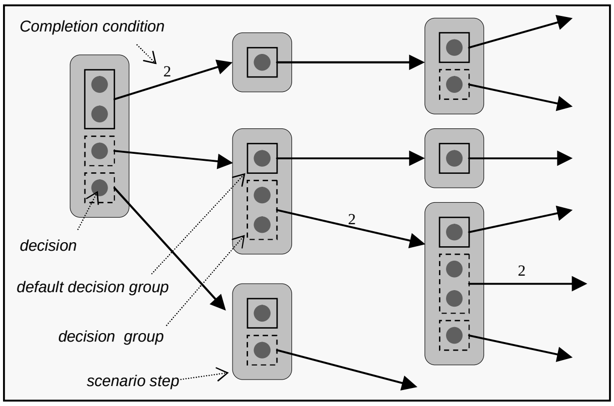


Figure 3. Interactive Scenario semantics: scenario steps, decision groups and decisions.

The INS itself is defined as a *scenario* (a decision tree) and a set of *profiles* which represents all the possible decisions a user can make (i.e. novice, expert). A *profile* is a list of references to the *decisions* of a *scenario*. Given a *profile*, the *scenario* will be limited to a sub-tree. Some inconsistency may occur in *scenario steps* that propose no *decision* allowed by the *profile*. In this case, the *default decision group* is triggered automatically (iterative selection of all *decisions*). As a side effect, an empty profile will provide a linear path along the defaults decisions (called “spectator” profile).

We defined a scenario description language for prototyping our Interactive Narration Spaces. Its pseudo-grammar is described below to give an overview of the INS semantics. Keywords are written in bold and non-terminal rules are in italics. The + represents the concatenation of terms (could be replaced by a comma but ‘+’ can be read as the “addition of information”). The $^+$ exponent stands for a 1..n repetition, while * stands for 0..n repetition. A statement enclosed in { } is optional and brackets are used for parameters.

INS ::= **scenario** (scenario_name) + *scenario* + *user_profile*$^+$
(INS is a scenario and a profile)

user_profile ::= **profile** (profile_name) + decision_name$^+$
(set of decisions allowed in the profile)

scenario ::= **root** (*scenario_step*) { + **options** (*decision*) }*
(root and the set of optional decisions)

scenario_step ::= **default** (*decision_group*) + *decision_group** {+ **options (***decision*) }*
(choices and the set of local optional decisions)

decision_group ::= *decision*+ + *completion_ratio* + **next** (*scenario_step*)
(set of decisions required to go to the next step)

decision ::= **decision** (decision_name) + *action* {+ *automatic_action* }
(action to perform and automatic action if timeout expires)

automatic_action ::= **auto** (*action* , timeout_value)
(the automatic action is performed after timeout_value time)

action ::= **action** (actionID) { + description }
(actionID used to communicate with runtime State Machine)

4.2. INS Authoring and Control

The Scenario Editor and Player (SEP) software is composed of three modes corresponding to each expertise requested for the design and usage of interactive scenarios: Declaration, Authoring and Execution (Figure 4).

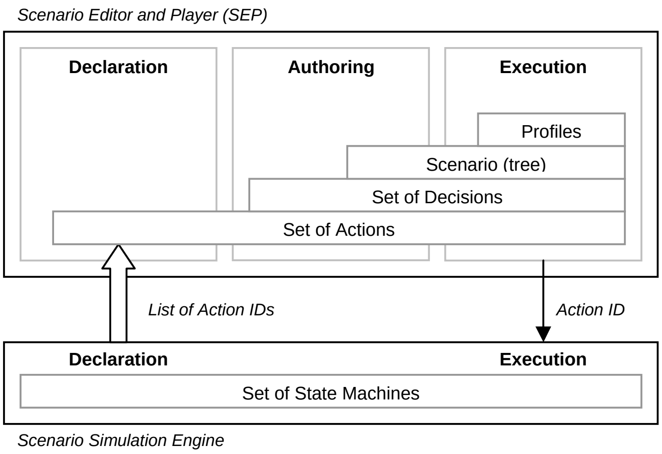


Figure 4. Scenario Editor and Player software architecture

At a first, *declarative* step of INS authoring process, the application domain is analyzed and the required content is produced: 3D scene, Virtual Humans, animations database, runtime engines, etc. On top of it, the *actions* are developed by programmers in the Scenario Execution Engine and are given a unique ID. An *action* is defined as a Finite State Machine (FSM) featuring the following possible states: *idle, activating, active, terminating*.

Once the pool of actions is provided, the application domain expert enters the authoring phase; the ‘story writer’, who does not need to be a computer specialist, uses the SEP tool to ‘write’ the interactive stories using the INS semantics. The Graphical User Interface constrains the creation but ensures the correctness of the scenario (see Annex 1: SEP GUI snapshots). Authors may then test the scenarios in a play mode.

The design of a scenario is not necessarily linear. As introduced by Dörner et al. [3], the authoring process requires incremental refinements of the scenario based on a good collaboration between storytellers and technology authors. The separation of SEP in three authoring tasks and the clear description of the semantics already provide a good basis for this communication. Moreover, as the scenario file built by SEP (*.sep) and the simulation engine only share the list of *actions*, their developments can evolve in parallel.

Although the current realization offers a good flexibility, the following issues remain open:

- Make a graph instead of a tree: we should develop a “GOTO” action to point to a labeled *scenario step*.
- Add milestone information: we implemented *State Variables* (a name, a value and a unit) that can be affected by *decisions*. They should now be re-injected in the scenario in order to affect the possibilities (e.g. with rules such as “Disable choice A if State Variable V is less than threshold L”).
- Specify the actors involved: we implemented *entities* (Virtual Humans or objects) that are relative to each *action*. They should now be used to sort the list of *decisions* by *entities*.

4.3. Social Channel interaction in INS

In order to implement the Social Channel interaction, we elaborated the concept of Virtual Assistant (VA) (a semi-autonomous agent). The VA is perceived by a trainee/patient as an inherent element of the evolving scenario. At the same time the VA's main role is to mediate the interactions between the INS and the subject.

In the course of the simulation, a trainee/patient (the decision maker) is accompanied by VA's. The user navigates, assesses the situation and makes *decisions* by issuing natural voice commands. VA's (decision executors) wait for commands and execute *actions* showing expected skills. A VA may refuse to execute commands that would push the scenario development into an undesired direction. In such situations, he may prompt the user for retrial or may suggest an alternative possibility. This way, all *steps* of the scenario offer a possibility of negotiation. In the extreme cases, when the participant is unable or refuses to decide, the VA can make automatically the *decisions* of the *default decision group*.

4.4. VR Simulation System Implementation

The immersive VR exposure system we developed is based on a wide-screen stereo display and a gesture based navigation. The Scenario Execution Engine is developed on top of the VHD++ framework (proprietary VR development meta-engine [12]) and drives a set of *actions* that are executed in parallel. They can support many simulation technologies (surround sound, object animation), but are centered on Virtual Humans animation (keyframes, procedural walk, animation blending, etc.), speech and behaviors (Python scripts).

The interaction between the user and the Virtual Environment is a complex issue that cannot be solved in a generic way. We considered that the interface module would be driven both by automatic solutions and by a human operator in order to overcome the common technological limitations of current interactive techniques (like voice recognition). In order to achieve this task, we ask an assistant (the Scenario Supervisor) to observe the simulation, interpret the user's inputs and act on the SEP software (*play* mode).

The system runs on 2 networked PCs under Windows 2000 operating system. One of the PCs, featuring Quadro4 900 XGL graphics board and 5.1 EAX compatible sound board, is responsible for immersive VR simulation. Depending on the simulation type, complexity of models, etc. the system can deliver on an average of ~15fps at ~100K Polygons scene load. The second workstation provides the trainer/therapist with the user interfaces of SEP allowing for authoring and control of the simulation executions.

5. CASE STUDY

5.1. A Decision-Training Application

In this section, we present the solution we implemented for the need of a medical European project dedicated to the training of paramedical personnel to emergency rescue [11].

The "BLS in the office" interactive scenario

The aim of this scenario is to train the *Basic Life Support* medical procedure. The trainee is immersed in a virtual office and discovers a man lying on the ground. He has to give BLS to the victim by giving orders to a young Virtual Assistant. The VA possesses all the skills required (mouth-to-mouth, chest compression) but highly hesitates on the procedure order; this is the job of the trainee. The latter will interact with the VA to apply his theoretical knowledge of the BLS procedure.

Table 1. Virtual Assistant – Trainee interactions during the BLS scenario

Virtual Assistant	Trainee expected response (BLS)
Come and explain the situation.	
"Hurry up, tell me what to do"	Check responsiveness
"He does not respond"	Open airways
"Good and now?"	Check breathing
"He does not breath"	Call for Help (phone 118)
Phone call, come back, "Good and now?"	Two effective breath
"Well, what's next?"	Assess circulation
"There is no pulse"	Compress chest
"And now, what should I do?"	Continue Rescue breathing & chest compression
Ambulance arrives	

Interactive story

This scenario consists of a dozen steps. At each step, the trainee has to make a (fast) decision. It does not present many possibilities because the BLS guidelines give no alternative. The scenario we implemented does not explore all the incorrect possibilities, which would all lead to a negative learning ('learn by errors'), but gives an example where the Social Channel can be used to orient the user in an Interactive Narration Space ('learn by success'). Table 1 gives a summary of the relation between the VA actions/speech and the response expected from the trainee.

Analysis

The main problems we encountered in the BLS scenario were related to the animation of the characters during the non-active phases, and to transitions between actions. The decision phase of the trainee may be long, and we should implement more *optional decisions* as well as add automatic idle behaviors.

The communication between SEP and the simulation engine should give more flexibility to the control of *actions*, and an improved State Machine model should be studied. A common management of *actions* states at runtime should offer a better blending/overlapping of animation.

5.2. A possible migration to an INS solution

In this part, we aim to analyze the need for psychotherapeutic exposures. For this purpose, we will use the 'Assertiveness' environment described by Klinger et al. [5] for the treatment of Social Phobia and compare the actual scene –based approach with an INS solution.

The clinical protocol they use for therapy involves two exposure phases: the first one is a free exploration of the environment, and the second one is guided by the therapist. Finally, the therapist shows "a model of behavior" for the encountered situation. The system was mainly tested with low immersion devices settings (keyboard, mouse and computer screen) that allows for a collaborative simulation. High-immersion devices (Head Mounted Display, mini-CAVE, etc) isolate the user, and an exterior intervention may hurt the Sense of Presence. Therefore, we argue that the therapist should help his patient indirectly by acting on the environment. One can imagine using different Virtual Assistants who can perform the 'model of behavior' on demand, so that the patient can observe how he could do, and try to imitate this action.

From another point of view, separating exposure in two phases (free/guided) should be synthesized in one single scenario in order to avoid the dichotomy 'what I do is bad' / 'what the therapist says is good'. As the 'Assertiveness' environment consists of a shopping situation where the user has to go out of a building to buy shoes, the storyboard could easily be translated to a linear path in a scenario tree, and other choices in the scenario steps should be introduced according to the dialog possibilities in each situations. All the simulation content has to be rearranged in action State Machines by a technical team.

6. CONCLUSION

We defined the needs for training and therapy and implemented an interactive scenario solution for building exposure. We found that the use of a Social Channel helps solving contradictions between narrative constraints (for trainer/therapist) and perceived interaction (for trainee/patient). We applied this system to training and are now waiting for validation results (currently performed by European partners). However, the procedure was satisfactory, and we aim to use this approach for psychotherapy. Our next objectives are to work with psychiatrists on the design of collaboration tools for the creation of therapeutic exposure scenarios, and to validate these experiments on Social Phobia treatment.

The idea of 'VR simulation with scenario' is well received by application domain specialists (as easily understandable and applicable) and can be kept as a base for exposures. Still the use of SC for interaction requires high level human-computer interfaces (voice recognition, speech synthesis, chatter bots, behavioral models, etc.) and our supervisor-based interface is of course not satisfactory in terms of autonomy of the system. The actual challenge is to progressively replace the supervisor with automatic software solutions, giving more and more autonomy to the Virtual Assistants, and offering higher-level control tools to the supervisors (guiding VA behaviors instead of decisions).

Another direction of research we are actually working on is the generalization of the Virtual Assistant concept to non-android Assistants. In general, any communicating entity, any 'intelligent' agent, regardless of its shape, could be considered as an Assistant if it provides help to the user. (e.g. distinguishing between driving a car, plugging an auto-pilot, and having a human driver). This way, the integration of any interaction in a scenario would be possible, but its naturalness would remains uncertain.

ACKNOWLEDGMENTS

Authors would like to thank Jan Ciger for his collaborative work on the concepts of Virtual Assistants, and all collaborators who helped for the proof reading of the manuscript.

REFERENCES

[1] A.H.Bullinger, A.Roessler, and F.Mueller-Spahn. From toy to tool: the development of immersive virtual reality environments for psychotherapy of specific phobias. *Stud Health Technol Inform*, 58:103-11, 1998.

[2] J.Burner, Towards a Theory of Instruction, New York, *WW Norton*, 1966

[3] R.Dörner, P.Grimm, D.F.Abawi. Synergies between Interactive Training Simulations and Digital Storytelling: a Component-based Framework In: *Computers & Graphics* 26 (2002), 1, pp. 45-55.

[4] S.Delp, P.Loan, C.Basdogan, J.Rosen, Surgical Simulation: An Emerging Technology for Training in Emergency Medicine, *Presence* vol 6 (1997) p147-159

[5] E.Klinger, I.Chemin, P.Legeron, S.Roy, F.Lauer and P.Nuges. Issues in the Design of Virtual Environments for the Treatment of Social Phobia. *First International Workshop on Virtual Reality Rehabilitation*, VRMHR 2002.

[6] R.Kuppersmith, R.Johnston, D.Moreau, R.Lofin, H.Jenkins. Building a Virtual Reality Temporal Bone Dissection Simulator, Medicine Meets Virtual Reality, *Studies in Health Technology and Informatics* 39 (1997)

[7] R.Lofin, P.Kenny, Training the Hubble Space Telescope Flight Team, *IEEE Computer Graphics and Applications* 15 (1995) p31-37.

[8] F. Mantovani, VR learning: Potential and Challenges for the Use of 3D Environments in Education and Training, Chapter 12 in *Towards Cyber-Psychology: Mind, Cognitions and Society in the Internet Age*, Amsterdam, IOS Press, © 2001, 2002, 2003.

[9] M.Macedonia, M.Zyda, D.Pratt, P.Barham, P.Zeswitz, NPSNET: A Networked Software Architecture for Large-Scale Virtual Environments, *Presence* vol 6, (1994) p265-287

[10] M.M.North, S.M.North, and J.R.Coble. Virtual reality therapy: an effective treatment for phobias. *Stud Health Technol Inform*, 58:112-9, 1998.

[11] M. Ponder ,B. Herbelin, T. Molet, S. Schertenlieb, B. Ulicny, G. Papagiannakis, N. Magnenat-Thalmann, D. Thalmann, Immersive VR Decision Training: Telling Interactive Stories Featuring Advanced Virtual Human Simulation Technologies, *9th Eurographics Workshop on Virtual Environments* (EGVE) 2003.

[12] M. Ponder, G. Papagiannakis, T. Molet, N. Magnenat-Thalmann, D. Thalmann, VHD++ Development Framework: Towards Extendible, Component Based VR/AR Simulation Engine Featuring Advanced Virtual Character Technologies, *Computer Graphics International* (CGI) 2003.

[13] D.P.Pertaub, M.Slater, and C.Barker. An experiment on fear of public speaking in virtual reality. *Stud Health Technol Inform*, 81: 372-8, 2001.

[14] B.O.Rothbaum and L.F. Hodges. The use of virtual reality exposure in the treatrnent of anxiety disorders. *Behav Modif* 23(4): 507-25, 1999.

[15] B.O.Rothbaum, L. Hodges, B.A. Watson, C.D. Kessler, and D. Opdyke. Virtual reality exposure therapy in the treatment of fear of flying: a case report. *Behav Res Ther*, 34(5-6): 477-81, 1996.

[16] R. Robb, Virtual Endoscopy: Evaluation Using the Visible Human Datasets and Comparison with Real Endoscopy in Patients, Medicine Meets Virtual Reality, *Studies in Health Technology and Informatics* 39 (1997) 195-206.

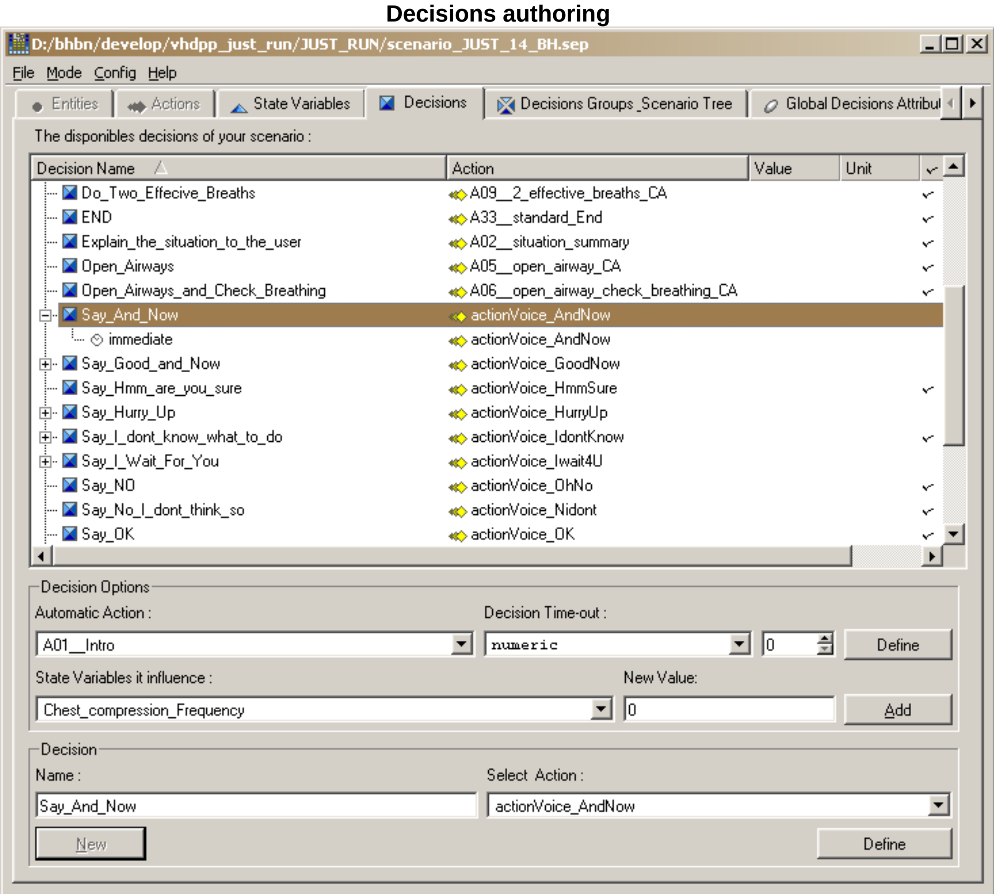
[17] F.Riquier, M.Stankovic, A.F.Chevalley. Virtual gazes for social exposure: Margot and Snow White. *First International Workshop on Virtual Reality Rehabilitation*, VRMHR 2002.

[18] Stanfield S, Shawver D, Sobel A, Prasad M, Tapia L. Design and Implementation of a Virtual Reality System and Its Application to Training Medical First Responders, *Presence*, vol 9(6), (2000) pp.524-556

[19] D.Tate,L.Silbert,T.King, Virtual Environments for Shipboard Firefighting Training,, *IEEE VR Conference*,1997

ANNEX 1: SEP GUI SNAPSHOTS

Decisions authoring



Scenario tree authoring (decisions groups in steps)