VIRTUAL PATIENT - INTERACTIVE STORYTELLING

H. Schulze

HZO Film & Medien
Liselotte-Herrmann-Str. 6, 10407 Berlin, Germany,
phone: +49/(0)172/7469394, mail: Heizo.Schulze@HZO-Film.de

Commission VI, WG VI/2

KEY WORDS: Virtual Patient, Interactive film, *QuickTime*, Type Movie, Data Dialog

ABSTRACT:

This e-learning simulation project makes it possible for medical students to practice conducting a diagnostic discussion with a "Virtual Patient". Based on these simulated conversations with a "Virtual Patient", the future doctor can practice diagnostic technique, trying to determine the ailment at hand. The patient is played by an actor who acts out a range of reactions. These responses are captured on digital video and organized in a databank. The fundamental principle is to analyze and represent the type of conversation that a doctor would have with his or her patient. Interaction occurs by entering text on the keyboard.

The project focuses on using cinematic means to immerse the user in a situation as close to a real conversation with a patient as possible. The basic idea behind the project is the non-medical specific use of human communication, interaction between individuals, which is essential. This could be expanded on, for example, to incorporate narrative. "Type-Movies" would be feasible, allowing the viewer to respond, in text form, to what is occurring in any space or even contacting the characters in a film, communicating directly with them.

1. INTRODUCTION

The combination of moving images - as one of the main technological basics of communication, and text - as the origin of human communication, is the concern of this work.

The topics in this paper vary from basic medical issues over communication theory to technical achievements. Elementary steps of patient-doctor interviewing practices are examined. A short overview about interactive film solutions at the last available state is given. They are objected in their various forms of interaction. Our main instinctive media knowledge is inspected, starting from cinematic rules.

A close look at the approaches of this work is summarized, which leads towards the elements of "Virtual Patient" from the didactic approach, the technical concept to steps of production and the single components and features in detail.

It has to be considered that the point of view of this presentation is that of a Designer and Filmmaker, not of a Scientist. Moreover of someone who is highly enthusiastic about the future of interactive film.

"Virtual Patient" is based on the research program "Visualization and Simulation Environments to Solve Difficult Learning Situations" [VASE1] of the "Learning Lab Lower Saxony" [L3S], Hannover. It has been developed to a preliminary state at the College of Fine Arts Braunschweig "Institute of Media Research" [IMF] in cooperation with the "Karolinska Institute", Stockholm [KI].

Fig. 1. "Virtual Patient" unfold in time

2. REVISED DIALOGUE TRAINING

The normal training situation for medical students of anamnesis diagnoses is a setting that

- won't bear repeating identical
- can't be reconsidered and comparable
- can't deliver a high emotional impact.

These elements are required to apply at the same time to deliver a measurable learning curve. Beside this, a presetting of a most realistic simulation of doctor-patient aural confrontation is required, before a real patient is met. A simulation to immerse the student into his future human-to-human confrontation is needed to focus on one of the main challenges the doctor has to solve: To gain information about the person, sitting in front of him - before the body is about to be examined.

3. PATIENT-DOCTOR INTERVIEWING PRACTICE AND BEYOND

Skills of special knowledge, e.g. "pathophysiology and clinical epidemiology, at one hand and knowledge of human behavior and social and cultural contexts" [Coulehan/Block, 2001] at the other are the two main basic fields required to interview a person. This can be done with different type of media.

Student - Book

The main source for our skills are books. From a book the student can retrieve the contextual meaning of different levels of human communication, such as "active listening skills, therapeutic core qualities, symptom words, respect, genuineness, emphaty" etc. and the different topics of "patient history" [Coulehan/Block, 2001] or facilitating skills of "questioning and relation-building" [Smith, 1996]. This is the basic information required, but it remains theory until the first interview takes place.

Student - Student

With a face to face practice students can apply their knowledge prior achieved. One student pretends to be ill - the other does the interview. They know each other and they are studying medicine, not drama. Here a first simulation takes place which helps to shift theory into practice.

Student - Actor

A practice with a trained actor with special skills, e.g. on patient history facts and symptoms, is another possible training method. This provides the highest emotional impact: the student is confronted with a real person. Although this is preferably the most realistic way to simulate a patient, it's far from being the efficient solution. There are strong limitations in availability, reproducibility, costs, reconsideration and comparability.

Student - Computer

Various e-learning solutions in all fields of knowledge are developed up to now. They all base on human-machine ergonomic schemes: the user interface with a combination of text, icons and active elements, such as buttons lead to forms of information presented as text, pictures, video, audio, 3D and combinations thereof. This environment is a simulation itself. As long as the main element (the simulated person) has not the topmost focus, such as in size, position, quality and realistic approach, the interaction still remains a human-machine connection.

Interactive Film

Definitions and solutions on the subject of interactivity in movies do vary from a branch or web-based structure and "spans to new narrative forms with unfamiliar names such as 'elastic media', 'Thinkies', and multithreaded interactive movies" [Hippolyte, 1995]. Any movie with a story time > presentation time can be declared as interactive film as a basic identification. This implies that the user/recipient has a choice to extend or shorten the presentation time. Besides the structure of nonlinear movies, the main challenge for interactivity in this media is the form of interactivity itself, the time interactivity can occur and the point of view (POV). The specifics of synthetic and real images/characters as well as the similarities and differences between interactive movies and games are significant as well. These topics will be dropped here to increase to focus on the most relevant items.

Form of Interaction

- a random selection, choice is not made by user
- context based selection by mouse click on icons or text
- visual selection by mouse click on elements in the movie
- language based interaction

Time of Interaction

- decision is made after segments of film
- decisions can be made while movie is playing within a preselections of choices (e.g. multiangle on DVDs, iTV, any level based media, such as games)
- decisions/choices all time

POV of Recipient

- 3rd person "they do"
  The characters are observed from a god like, objective position (master shot).
- 2nd person "you do"
  The recipient identifies and acts together with the protagonist (e.g. over shoulder shot).
- 1st person "I do"
  The recipient takes part in the story from a subjective position
  (position of camera = position of protagonist AND protagonist = recipient).

Examples

Most of the interactive media combine elements within the given categories. The following list determines the essential elements in each category. The examples are chosen to present the variety of media which has to be inspected to consider useful elements and forms of interaction.

"Myst, Riven, Exile" © Cyan Worlds, Inc.,
http://www.myst.com, adventure games, 3D environment,
Form: visual, Time: preselection, POV: 1st

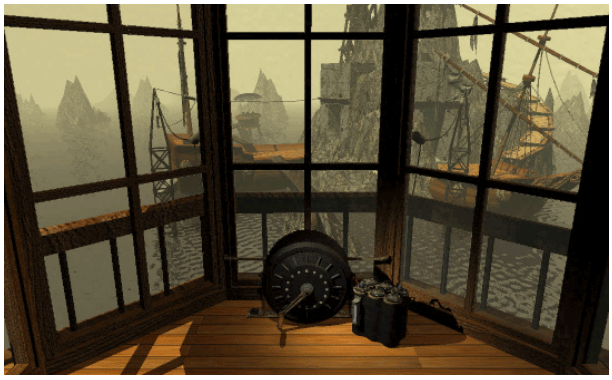

Fig. 2. The adventure game "Myst" © Cyan Worlds, Inc.

"Virtual Cinema" © Hyper Bole Studios,
http://www.hyperbole.com ,
authoring software for interactive movies,
Form: context based, Time: after,
POV: 2nd/1st person

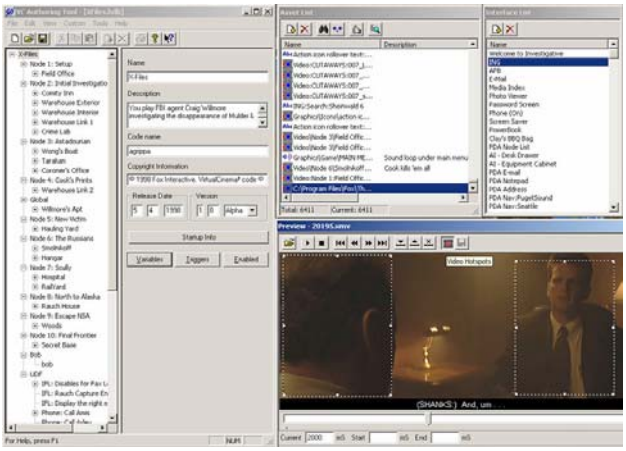

Fig. 3. The Authoring Software "Virtual Cinema" © Hyper Bole Studios

"Korsakow" © Florian Thalhofer, http://www.korsakow.com,
authoring software for keyword based interactive movies,
Form: visual, Time: after, POV: any

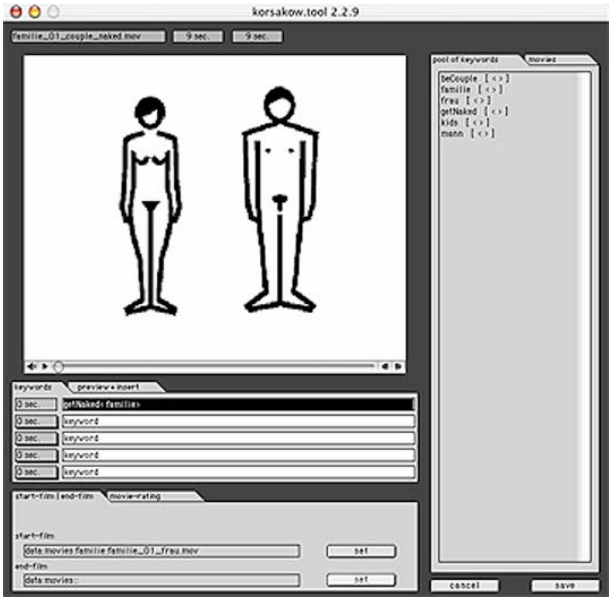

Fig. 4. The Authoring Software "Korsakow" © Florian Thalhofer

"Subservient Chicken" © Burger King,
http://www.subservientchicken.com,
online advertisement as interactive movie, Form: language,
Time: all time,
POV: 1st person

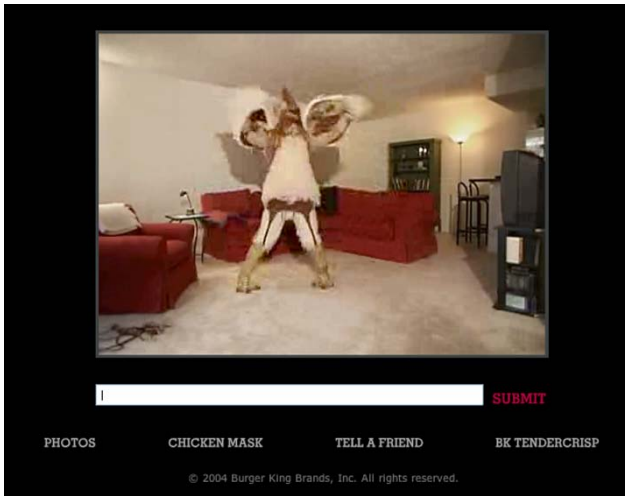

Fig. 5. "Subservient Chicken" online advertisement as interactive movie © Burger King

"ISP Medical History Module" (ISP-VL)" © Karolinska Institute Stockholm,
http://www.lime.ki.se
patient interview simulation,
Form: language, Time: after,
POV: 1st person

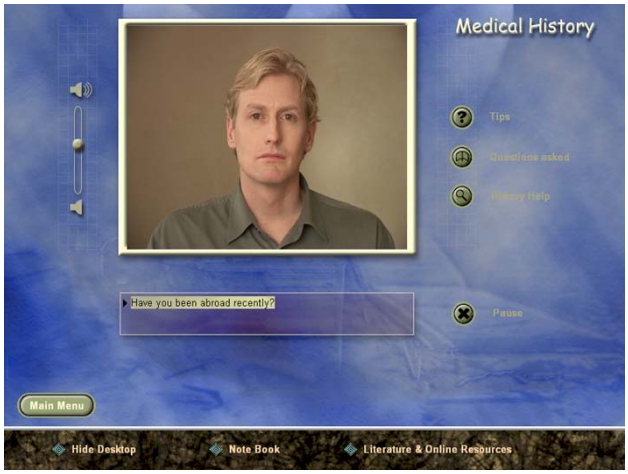

Fig. 6. "ISP Medical History Module" © Karolinska Institute

4. A VIRTUAL PATIENT IN INTERACTIVE FILM

4.1 Interfere with Language

The written word is currently experiencing a renaissance. SMS, email, and chat have "thrown us back" upon our original form of indirect human communication. Yet we automatically associate interaction with the man-machine link, with the computer as a tool that extends the human arm and the symbiosis this produces.

Let us now image we could actively and interactively participate in a film. And our only means of intervening is to flail away with the extension of our arm (mouse or joystick) and to "poke around" with every click: In a manner far removed from how we are used to viewing films, we poke buttons, houses, cars, plants, and people. I feel that current approaches of this sort can at best be termed active because they primarily emphasize motor skills.

What if we really could intervene? Intervene? Yes, I mean becoming interactive in the sense that we could add our opinion to the action! Yet to do so requires that we first downgrade from all the 3D game graphics to the medium I am using here to communicate with you: letters of the alphabet. I would like to intervene in the interactive film by including the keyboard's primary function (M is M and not "shield" and the space bar is a space and not "fire"). I can "speak" to the figures in my film, maybe in a dialog, as an inner voice, as almighty God, or I speak for them.

My research of interactive film is based to a great extend on the experience I have acquired in the international research project "Virtual Patient". This experience involves the semantic structuring and synthesis of a simulated learning environment that allows communication between the user and video material that has been produced in advance, thus enabling the trainee to experience the content from within the context of the film.

4.2 Student - Interactive Film

"Virtual Patient" is studying the possibility of activating, motivating and creating an emotional relationship between students and the characters in simulated learning systems. This is achieved by studying the impact of different techniques to simulate patient history interviewing.

A simulated dialog between physician and patient gives medical students the opportunity to test their diagnostic skills. They can use the keyboard to ask the virtual patient questions that the patient then immediately answers. The application primarily utilizes the language of film to achieve the flow of subjective experience. The basis for the application consists of an extendable system of keyword links that uses a special search algorithm in XML data, a database with specific and generic answers in the form of film clips, and a programmed algorithm for dynamic real-time editing with QuickTime.

Since the form of interaction has to be language based, the time of interaction has to be continuously and the POV must be 1st person for a subjective immersion.

This project continued where the existing WGLN VSP-VL [ISP-VL] and VASE 1 [VASE1] project left off, utilizing recent developments in full-motion video simulation based on editing patterns established in cinema, in real time. Modern cinematic techniques make a greater degree of realism and depth possible. A main demand for further application has been the aim to use common software. This is required to extend the area of possible users, which run ordinary computers with no hardware extensions. The second attention has been the hybrid employment to develop one solution for local and online use. The only difference of both versions should be video quality, because of bandwidth limits. The online version should still offers all other features.

The list of required features range from:

- large scaled video
- local and online use
- platform independent
- realtime edit
- dynamic montage
- extendibility
- maximum of immersion
- to ease of use

5. THE PARSING AND OTHER APPROACHES

5.1 Text Interaction

As we eliminate other interaction elements other than typed language, there is a strong focus on the keyword parsing. Basically one-word-questions have to be prevented and there are additional elements to provide a satisfactory result. E.g. short term memory keywords have to be read with each answer and dropped with additional answers, regional different meanings might be available, sensitive information is triggered with a continuous questioning only and the character might respond in different moods.

5.2 Video Masses

Because of the cinematic focus of the project, answers are shot with three cameras simultaneously (to provide movie montage behavior). This has to be solved in storage and steps of production and documentation.

5.3 Dynamic Editing and Realtime Behavior

The most challenging feature is the request for a continuous movie-like experience. Together with large scaled video, even in local use this request has to solve various obstacles, such as speed in parsing and loading, as well as the problem to load and play a movie without a noticeable stop or glitch. Not to mention the online problems.

6. HOW TO BRING LIFE TO THE IMAGES?

6.1 Didactic Approach

The didactic approach focuses on improving and encouraging communication in the simulated "doctor/patient" situation. There are clear stages in this type of conversation: the introduction, small talk to "break the ice", the diagnostic issues at hand, waiting (the patient waits for reactions at various points in a diagnostic conversation), comments (to the doctor's remarks, not questions per se) and various reasons for breakdown in communication (misspelled questions, time out,

loosing the patient's confidence). This ought to aid the process of establishing some sort of relationship.

6.2 Technical Concept

We looked for a widely accepted software which is able to cope with all our required tasks and configurations. A software that has been sufficiently developed and is easy to install. We had to learn more about QuickTime software of Apple Inc., to take full advantage of its possibilities. In this cross platform solution, which works as a digital container of various communication media in the virtual digital world, one can manipulate and edit the required movies ad hoc. We investigated the use of codex on larger scale video images and developed a scripting solution using (the programming language) Qscript for editing real-time events.

7. THE VIRTUAL PATIENT

7.1 Working Procedure

The concept can be divided into four sub-topics: the visual/aesthetic framework, the human interaction/parsing, production and authoring/programming.

7.2 Defining Aesthetic Elements (Cinematic)

The images provide the visual and logical framework for the usage of interactive video. The decision was taken to "immerse" the viewer, creating a simulated patient. It is then imperative to a) capture the user's attention in a film-like manner and b) increase interactivity. As a result, the following possibilities emerge: a) a dynamic montage of the patient's response, or b) interaction within virtual space/film invoking the patient's response to the doctor's physical presence. Altering the frame size and composition in realtime can help maintain motivation and provide a immersive film experience.

Fig. 7. Shooting Impression

7.3 Classification of Human Interaction

The starting point has been a database with the main medical related topics in patient-doctor interviewing, based on the medical experience of the project partners at the Karolinska Institute. This database has been altered and extended with major elements in a broad range of human reactions, such as general answers, red herrings and personal details of a predefined character.

The parsing is and will be based on keywords which yet provide a specific response. Three things which will improve the results to the patients response: 1) a keyword area which will help to prevent unwanted or illogical answer from being formulated, 2) the further development, by monitoring recent questions/ answers to assure that all reactions reflect what the patients require, 3) the addition of communication headings and sub-divisions to retain the structure and yet allow it to expand as is needed. Answers can be added continuously.

7.4 Pre-, Production and Post-Production

Main steps of movie production where required to:

- write a script and a description of the character
- cast a actor
- build a team
- rent a studio
- build the set
- maintain technical equipment
- shoot within a week
- capture and edit material
- compress files
- maintain database entries

Fig. 8. Shooting in Studio

8. AUTHORING INTERACTIVE QUICKTIME

The logic structure is implemented within a interactive QuickTime, also known as wired QuickTime. Through the complex, but efficient programming language "QScript" used in the QuickTime authoring software "LivestagePro", various algorithms have been developed, which offer the required results. Although a missing support in parsing delayed the whole project, which actually terminated end of 2003. With the development of a 3rd party component "QSXE", which extends massive the wired abilities of QuickTime, the further development of the project has been done on the own behalf of Heizo Schulze. The result has been a complete new programming to implement the features of QSXE within 2004.

8.1 How to Prevent the Loop?

One main element to detain a repetitive experience are the available frame sizes of each clip (A - Medium, B - Close Up, C- Extreme Close Up). Following cinematic basic rules a "A-

clip" can follow B or C, but not A. "B-clip" and "C-clip" are treated similar: no clip of the same size will be loaded after one segment has finished.

Fig. 9. Three Frame Sizes of each Clip for dynamic Editing

Additional situations of interactivity states are defined - which will be played for specific use. Such as: Intro, Waiting, Repetition, Misunderstood, Specify, Timeout, Interruption, Insulting and Leaving. These are called General Responses.

The key feature is the classification and link of three objects: Parsing, Answers and Clips. Keywords are NOT linked to static clips, they are linked to Answers. Each Answer can have associated clips for Neutral, Hesitant, Aggressive response, can hold up to 3 Repetitions and up to 3 Follow Ups, 1 Short Time Memory and currently not used triggers such as: EU/US Region, Question Required and Sensitive Topic.

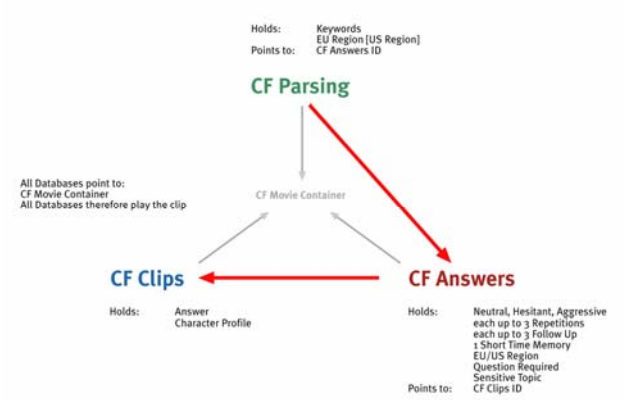

Fig. 10. Parsing - Answers - Clip Connectivity in Database and XML

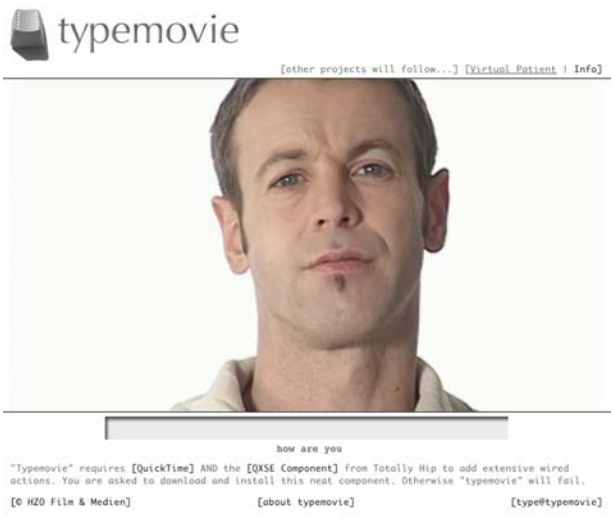

Fig. 11. The current state of this "Virtual Patient" project and additional text interaction based movies can hopefully be found in near future at: http://www.typemovie.com

8.2 Storyworlds

I consider the examination of the following scenarios for further narratives:
User Actor (one way communication)
Question Answer

User Actor (alternating directions)
Question Answer
Answer Question

User Scene (one way, first person, 2nd person or 3rd person)
Question Answer

User Scene (alternating, first person, 2nd person or 3rd person)
Question Answer
Answer Question

8.3 Cinematic and Historical Examples

As one example how to use text as interaction in narratives I would like to take this opportunity to apply in theory the use of "Type-Movie" technology, to relay ideas and make the cinematic exploration of history possible.
Most of what our forefathers created is now perceived as lifeless, even dull material. Castles, museums, monasteries and cloisters, prisons, ruins, even factories are all buildings with some semblance of complete interiors, and they also contain historic objects. When we enter them we are immediately immersed in their atmosphere. They impress, yet remain distant, regardless of how much time we spend there. They are loaded with historical, human emotion. It is only when we begin to learn about the previous residents, who created connections of all sorts, and that is precisely that which gets us emotionally involved, and we cease being outsiders and approach what was once foreign, or even feel that we understand it.

There are, basically, two different approaches that one can take:

a) a mosaic-like structure, or b)narrative form.

1) relies on the collection of information using codes, and the decoding of these codes proceeds, based, for example, on text, images, or architecture. Only those trained in the field are able to assemble this mosaic, as it forms a counterpole to pure authenticity.

2) on the other hand, is based on emotional response, i.e. in the form of a book, a film, or a very vivid presentation delivered by a gifted and motivated art historian while conducting a tour.

Current interactive media, unfortunately, function according to a). Regardless of how complex and extensive they may be. Using "TalkMovie", I would aim for a synthesis of a) and b). Self-contained spaces form the framework for a). This is, undoubtedly, a conventional notion. What adds a new dimension to all of this, beyond conventional exploration of space, would be the introduction of life to these spaces.

In one small room in the cloisters, for example, one would actually find a nun going about her everyday routine, hundreds of years ago. Mother Superior has taken her special place, at the fireside, and, in the kitchen, dinner is being prepared. Using the keyboard, I can, at any moment, establish contact with them or ask questions. They respond, answer, or show me something and I learn a bit about their lives, in a world hundreds of years ago.

Each separate room remains an island unto itself. The user has the option, due to the complexity involved, of a limited extent of immersion in the convent's activities. Each "station" on one's visit would form a stone in the mosaic. One could even consider using a figure/character to explain elements of the societal and political issues of the period. This would enable the human-friendly transfer of information to occur in a new manner.

9. DIALOGUE OF DATA

QSXE has to be installed in order to take advantage of "Virtual Patient".

The Master Movie - the Data Shell
The Movie itself is about 52KB. It holds all algorithms for parsing, dynamic processing etc.

QuickTime Movies
Over 1250 clips have been shot with 3 cameras simultaneously (Medium, Close Up, Extreme Close Up).

Filemaker Database
Filemaker is used to incorporate the various connections between clips, parsing and answers.

Fig. 12. Database Clips

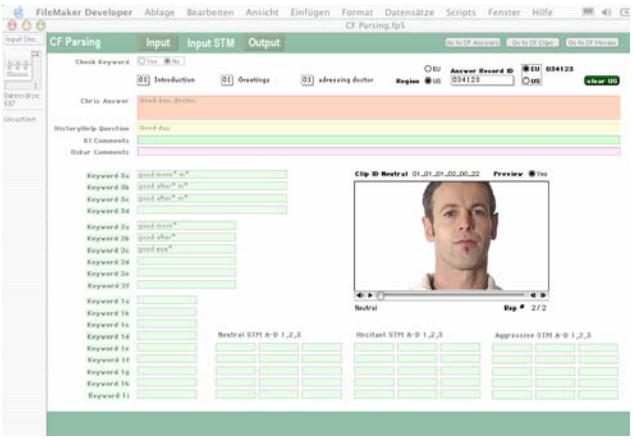

Fig. 13. Database Parsing

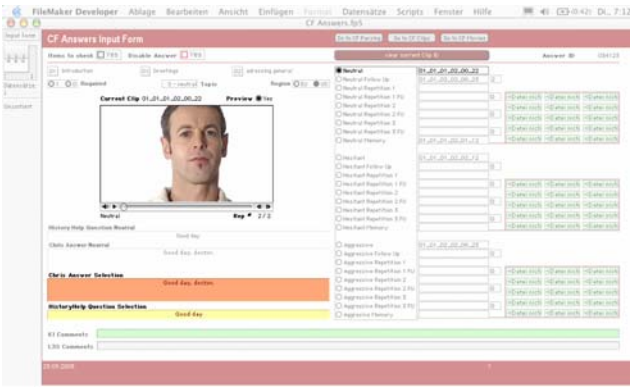

Fig. 14. Database Answers

XML Structure
XML files are the essential knowledge of the master movie and are loaded on startup. These files hold all information of the database setups.

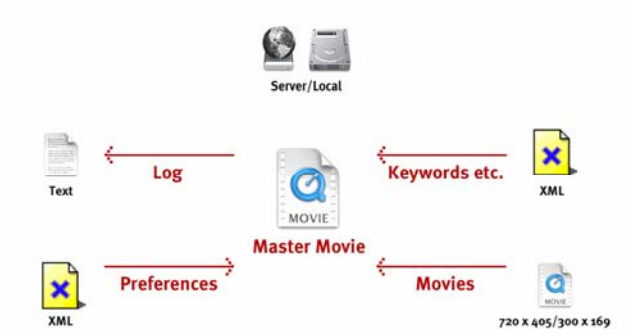

Fig. 15. Dataflow

Log Files
User input is monitored and stored remotely, such as answer, keywords found, played clip, etc..

```
type: VirtualPatient-Log/Apple_4151_6.520_1_15.1.2005 _No _-
23295_Time_22-22-0_Session_1.txt
Question: hello
Answer: Hello, nice to see you, doctor.
AnswerID:  34129
Clip: 01_01_01_02_00_12
KeywordString:  hi hello
KeywordMatches:  34129 :  hi hello
UserTime: 0.546 Minutes after last input - at: 22:22:24
```

Fig. 16. Log File Example

10. CONCLUSION

The different frame sizes with the realtime edit and the dynamic montage as the major part evoke the aimed emotional involvement of the user. A big emotional impact and immersion has been obtained. There have been doubts about the technical feasibility regarding the online solution. The results where able to wipe this uncertainty. The current state of "Virtual Patient" shows a huge potential for extensive use in dialogue simulations of various area - local and/or online.

There are further steps for development to follow soon: extension of the answers, clips and keyword databases, therefore the addition of available reactions and answers of the "Virtual Patient". In near future a linguist should provide assistance for the optimization of the parsing process on the keyword basis, such as short term memory keywords, regional differences, repetitions, topics with sensitive information and mood changes of the character. These elements are already incorporated in the interactive algorithms, but are not yet used in the database.

Speech recognition seems to be a possible combination/extension to eliminate the restrictions of human-computer interaction. In addition the depicted character needs further attention in terms of comparability between a "real" person and a synthetic image, such as 2D or 3D animated characters. The extension of additional reactions/answers with these types of visuals would offer a broader and flexible feasibility, but may lack in immersion of the user.

11. REFERENCES

typemovie, "Virtual Patient" online, http://www.typemovie.com/

HZO-Film & Medien, Portfolio of Heizo Schulze, http://www.HZO-Film.de/

[L3S] Learning Lab Lower Saxony, Hannover, http://www.l3s.de/

[VASE1] Visualization and Simulation Environments to Solve Difficult Learning Situations, L3S, http://www.l3s.de/vase1/

[KI] Karolinska Institute, Stockholm, http://www.lime.ki.se/

[IMF] Institute of Media Research, Braunschweig, http://www.hbk-bs.de/

[ISP-VL] Interactive Simulation of Patients, KI, http://www.lime.ki.se/cul_cd_et_projects_isp.htm#226

[Coulehan/Block, 2001] Coulehan, John L., Block, Marian R., "The Medical Interview", F.A. Davis Company, Philadelphia, 2001

[Smith, 1996] Smith, Robert C., "The Patient's Story", Little, Brown and Company, Boston, 1996

[Hippolyte, 1995] Hippolyte, Michael St., "Movies of the Future: Storytelling with Computers", 1995

12. ACKNOWLEDGMENTS

It goes without saying that my research fellows deserves many thanks: Rolf Bergin, Uno Fors, Olivier Courteille from KI. Uli Plank for supporting my work from the beginning of my time at the IMF. Christoph Strutz for doing so at the L3S. Kathrin Koch for her tireless emotional support.