Beyond Virtual Tutors – Semi-Autonomous Characters as Learning Companions

Ulrike Spierling
Fachhochschule Erfurt,
University of Applied Sciences
spierling@fh-erfurt.de

Abstract

How does an electronic interactive learning medium support constructive learning, beyond mere instructions and the supply of predefined exercises? This will be discussed in this paper, along with the consideration of new technical possibilities of virtual interoperable characters as learning companions. Following a brief summary of state-of-the-art Interactive Storytelling issues to be considered for the creation process, first experiences from two projects with different approaches are described that employ playful text-based conversations with graphical animated bots. The results lead to the (surprising) conclusion that the most significant learning process seemed to be achieved for the active authors of the conversational dialogues. Beyond an instructive "virtual tutor", primarily misunderstood as replacing a human educator, different metaphors are suggested that change some expectations for learning with virtual characters. In the consequence, learners will be viewed as authors, who create their own virtual companion and shape it on their own responsibility.

1 Introduction and Motivation

The concept of Interactive Digital Storytelling has the potential to become a paradigm for future interactive knowledge media. The basic idea couples dramatic narrative with the interactions of users, providing the highest forms of engagement and immersion. It also stands for the connection of games and stories by utilizing inherent structural elements of both. For teaching, storytelling is used to communicate complex information through the reference to human problems, mapped to a set of characters with a plot. In doing so, factual information is often represented by text or spoken dialogue of characters. Gaming is of use not only for entertainment, but also for the training of skills and for constructive learning in a simulated scenario, serving as a fault forgiving environment. Compared to stories, games provide little explicit knowledge transfer, compared to their high potential for acquiring tacit knowledge.

Digital storytelling agents can achieve more than simply being virtual guides or tutors, which are commonplace today in a variety of software products and concepts. As in stories on a stage, their role could be to interact with each other as a set of characters to present a dramatic storyline; and as in games, they have the potential to serve as all sorts of sparring partners for players to interact with, such as representing the bad guys, or companions who ask for help. This exceeds the view of a virtual character, staring at a user in order to help out with the interface, or "to be the interface" for something even more technical.

Figure 1 shows a scheme of 3 different kinds of human interaction concepts with new media, which are present at the same time when interacting with media applications, however possibly with changing biases in their relevance:

1) Mediated communication, Human - Human
2) Factual experienced actions, Human - Media
3) Medium for thinking, "extension of man", Human - Self

All three concepts can be considered for learning and knowledge media: Storytelling and teaching can be identified as the Human to Human concept, no matter if a teacher makes a chalk drawing, or if she scripts animated talking characters in order to convey information. According to the third concept, a sheet of paper can serve as an extension of a Human's own imagination or memory in order to think and sketch hypotheses, and so do simulations of virtual characters, leading to an evaluation of hypothetical extrapolations. Finally, in connection with both, a learning effect also can arise from the factual way of interaction with the medium itself; so do writing exercises lead to better writing, and drawing and speaking train to draw and to speak.

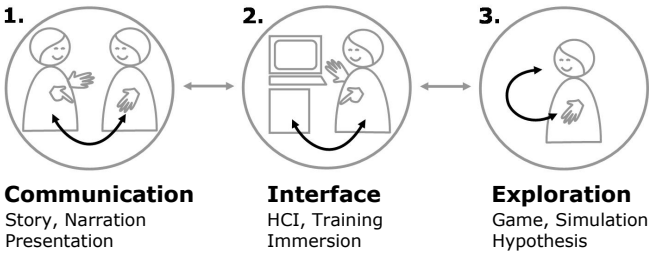

Figure 1. Human-media interaction concepts

Within the context of this paper, learning scenarios and tools have been envisioned, which allow the training of verbal dialogues and the simulation of conversational behavior. In combining the aspects of stories and games, factual information can be conveyed by authored spoken text of virtual agents integrated in an audiovisual presentation, and different opinions can be rendered by mapping them on several agents in a dispute. Finally, verbal interactions allow users to test their own decisions and participate by active construction of a dialogue. This paper reports on experiments with two technical platforms that have both been tested in educational environments at university level.

2 State of the Art and Related Work

The technical combination of gaming and storytelling raises a number of conflicts, which recently have been discussed in the computer games community at length, e.g. [Frasca 2003]. In addition, successful implementations providing intelligent conversations with animated virtual characters are rare, and there is no real success on the entertainment market to date. The reason is seen in the dilemma that in order to be dramatic, a story needs an immensely higher complexity in its structure than a game. In a novel or movie, it is tuned and laid out once in a linear order by human authors. For adding interactivity and not losing dramatic tension, AI techniques need to be employed to generate suitable plot developments in reaction to participants' free actions. Seminal research in this field has been conducted in the early nineties; for a number of references see the OZ project webpage [OZ 2002]. The result is an emergent behavior of intelligent agents controlled by a drama manager. Since AI techniques are hard to handle by creative writers and educators, recent research

projects examine middle courses between autonomous characters providing emergent interactions, and a more author-friendly plot driven approach.

A major example is "Façade", developed by M. Mateas and A. Stern, which is the first application with natural language conversations between two characters and a user [Mateas and Stern 2003]. Façade allows users to participate in a predefined and pre-recorded dialogue between virtual characters and to cause them to change the result of their dispute at certain points, while the agents do not react to every input. Mateas and Stern implemented their own dialogue management system, which is currently not aimed at being authored by non-programmers. In particular, the only usage scenario is targeted at the end-user, and dialogues can not be created on the fly.

The project "art-E-fact" [Spierling and Iurgel 2003] presents a similar integration of simulation and plot. In contrast to Façade, an authoring system is central to the way a story is built in art-E-fact. First, a synthetic text-to-speech system with lip-sync animation enables immediate audiovisual playback and interaction with created dialogue elements. Further, story writers are able to transfer a certain degree of factual knowledge through dialogue, using a hierarchical and modular control approach. They start with a graph of concrete dialogue acts, similar to branching, and provide interactions by adding rules within detailed nodes of the graph. The dialogue acts refer to the idea of speech act theory [Searle 1975], and support the definition of semantic and non-verbal statements. The art-E-fact approach supports easy story creation with predictable user interactions. However, more complex and free conversations are difficult to handle.

A project with a similar goal, but with an approach from the opposite end as art-E-fact, is "Scenejo" [Müller et. al. 2005]. From the start, AIML pattern matching [ALICE 2005] is used to provide free conversational interaction with users, which can also be supported by modeling dialogue acts on a basic level. Later, a story graph allows writers to line up conversational scenes and their parameters. Scenejo is an experimental improvisation stage, in that it provides a playful test environment for emergent conversations between chatbots that are hard to anticipate.

In order to interact with believable representations of the virtual characters through multimodal techniques, the research field of "Embodied Conversational Agents" also provides major contributions to the relevant state of the art. As an example see [Cassell and Bickmore 1999]. These trends rather point towards the vision of creating a fully-fledged Virtual Human, than towards solving authoring problems for interactivity.

In summary, the efforts done in modeling semi-autonomous virtual characters and their dialogues still do not meet the needs of writers tackling the dilemma of interactive storytelling. Moreover, writers as well as educators have difficulty articulating their relevant needs as precise as necessary to formalize them. Since the telling of a story with a message is a human endeavor, we need more practical experiments with existing platform prototypes, providing step-by-step insights to the processes. In the remainder of this paper, no new technical platform is introduced, but examples of such experiments are reported, here with the art-E-fact and the Scenejo platform.

In the end, and simply non-technically, the educational ideas presented in the introduction have already been applied for years within the community of Gaming Simulation, striving for constructivist methods of learning [ISAGA 2004]. Complex systems (such as an ecosystem or a social system) can be modeled with a reduced set of parameters and rules of their interdependence. A learner can take a role, step into the simplified system, make decisions or invoke events and directly experience results during the role play. Roles can easily be taken by human co-learners. These direct experiences only become meaningful as learned bits during a "debriefing" phase with a human educator, after the role has been abandoned and a more neutral view on the systemic procedures—including own decision making—can be obtained. As a conclusion, the vision of a possible replacement of a teacher by a virtual character is fading away, not only because of technical problems. Using virtual characters as co-learners, sparring partners or companions is a more compelling metaphor, because it supports constructivism. Additionally, we may start out and create something meaningful even with a simple set of parameters.

3 Experimental Findings

The need for practical experiments with the mentioned authoring platforms art-E-fact and Scenejo was the motivation to first start an interdisciplinary cooperation of students from two German universities in the fields of applied computer science and education science, aiming at the development of new forms of e-learning for elementary school. The goal for students from mathematical didactics was to rethink and develop strategies involving interactive text-based dialogues with virtual characters in the context of a learning concept. The undergraduate computer science students first aimed at quickly acquiring sufficient knowledge for the implementation of such interactive dialogues in a graphical environment. This included concepts of Web3D (we used VRML/X3D), speech synthesis, chatbot authoring (we used AIML), character design, and state machines. In a second course only with the computer scientists, the goal was authoring of characterized chatbot knowledge bases (with AIML) to study emergent conversations within the Scenejo platform.

First of all, each of the design processes pointed to the need for suitable conceptual models in helping the authoring novices also grasp the general purpose of Interactive Digital Storytelling. Therefore, abstract design concepts of interactive storytelling were introduced through their supervisors, including character design, story design and planning, as well as interaction design.

After the first half of the semester, project teams were set up where each team consisted of students from both faculties. The goal of the team work was to conceive a small learning environment for mathematics, utilizing spoken language and dialogues as the central means of interaction. In a game-like and entertaining graphical world (compare Figure 2), the conversation with virtual actors should show children how to "speak" about mathematical problems, to find new ways to solve these problems – not just to present a solution – and to explore mathematical ideas from multiple perspectives.

Figure 2. 3D example scene with interacting characters

The different teams independently created examples for a learning environment based on storytelling or game aspects, each one for a different didactic problem in mathematics. The resulting story worlds for mathematical conversations included examples such as a magic forest from which one has to escape by solving arithmetical tasks (compare Figure 2), and scenes inside a molehill with a similar break-out mission (see Figure 3). Suggested through the art-E-fact platform, a basic scheme was to use at least two characters performing staged dialogues with each other, and prompt the user for action upon the author's request.

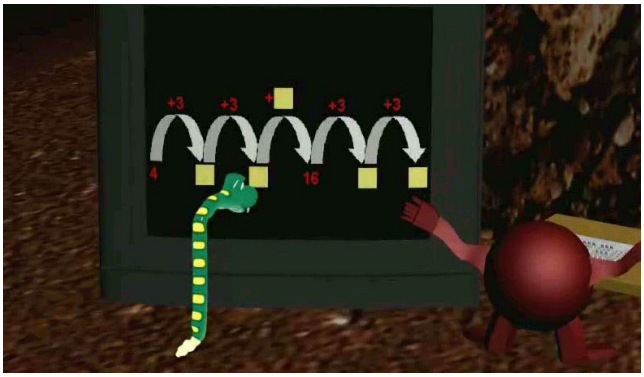

Figure 3. Student's example: Two characters (the Little Scary Mole and the Worm Helper) present numerical challenges

As an applied principle of storytelling, at least two characters are needed in order to articulate their personality, preferably with complex interpersonal relationships. For an e-learning application for children, these relationships initially are abstract roles such as a teacher, a co-learner, or a learner, represented dramatically for instance in scenarios such as "the smart and the dumb", "the sorcerer and the apprentice", "the villain and the helper" etc. These combinations induce much of the fun and interesting dialogues, and also assign a role for the interacting learner. Mostly, well designed characters live in an accordingly designed world that matches their goals and personality. This world can illustrate a very special and unusual situation by transferring it into a fantasy setting.

The authoring of the stories was performed at different levels. On the scene description level, all elements of the scenes in a story are represented as VRML-97/X3D objects, and virtual characters were modeled corresponding to the H-Anim standard. On the character design level, individual character states are modeled in terms of behavior and their emotional attitudes. In art-E-fact, this primarily contains the description of basic movements and gestures, as well as other characteristics such as speech sound. Further characterization becomes evident through the course of the characters' dialogue lines.

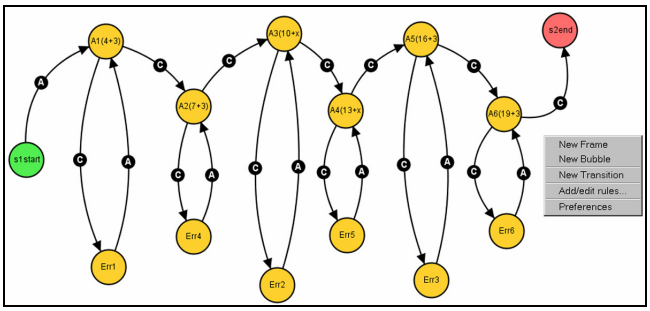

Figure 4. Transition diagrams revealed (un-)favourable dialogue strategies for question-answer situations

On this level, the conversation can be modeled in terms of a transition network similar to a branching structure (compare Figure 4). The authoring interface provides a graph-like view to the dialogue structure, where dialogue acts of virtual characters are represented as circular nodes. The edges in the graph represent transitions between them, which can be connected to specific conditions. Possible conditions mainly involve a text pattern matching functionality, supported by AIML expressions, as well as time-out function definition. While planning the interactive dialogues, educationalists had to anticipate any possible answer of a child, and they had to come up with various potential virtual characters' reactions.

In contrast to this approach, in the Scenejo platform a course of dialogical interaction can only be modeled through the addition of matching text patterns to a specific search-space, forming a knowledge base of a bot. The result is a less predictable, emerging dialogue, influenced by the users' input and the bots' databases of possible lines matching a text pattern coming from either a user or another bot character. Figure 5 shows an example screen of the current prototype, which allows for setting up various numbers of characters on-the-fly, and test their interaction without having to explicitly model a dialogue graph.

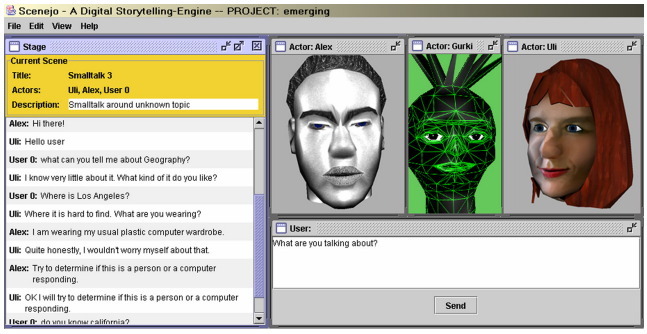

Figure 5. Emergent dialogues in Scenejo, resulting in a competitive small talk

The bigger the search domain, the more interesting the dialogue will be, yet this comes at the cost of becoming unpredictable. Therefore, also Scenejo provides a story graph structure, which partitions the search spaces in order to provide control over the global conversational plot (compare Figure 6). The difference to the art-E-fact graph structure is the level of detail where it is applied.

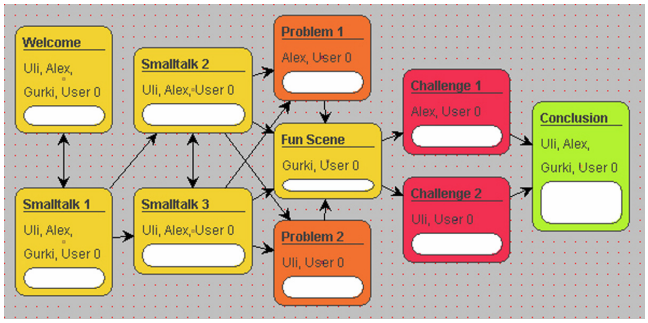

Figure 6. The Scenejo story graph editor

With Scenejo, it turned out to be more difficult to model the conceived dialogues, without having a background in programming. In contrast to that, the art-E-fact graph tool put forth remarkable realizations of conversational concepts. The structure that graphically emerged, similar to a navigational map through a dialogue, revealed strategies for dealing with specific

situations in a learning scenario. For instance, in cases when almost any contribution of the interacting child would be a wrong answer to an arithmetic problem of the mathematical scenario, turning back again and again to the same question results in a tiresome and frustrating reiteration (compare Figure 4). In a real classroom, elementary teachers would solve these situations by making use of analogies, comparing, and giving open questions preference over closed ones. Based on their analysis, all teams came up with preferable dialogue structures for the given learning scenario. As a side effect, this provided the students with deeper insight on didactic scenarios and their problems. Since up to this experiment none of them would have used a similar visualization to compare their dialogue strategies during lessons, they suggested analysing their classes by means of the graph images in order to uncover their real-life interaction structures.

While all teams produced complex and very interesting results within one semester, they did not succeed in developing a complete solution for an educational application in their context. However, during the design phase, major findings have been derived concerning the potentials of Interactive Digital Storytelling applications of this kind. In drawing a comparison between the two approaches, the use of the transition graph approach to dialogue modelling yielded a real and direct learning effect – concerning insights into the didactical dialogue structures of the educationalists. The modelling of a search domain instead proved to result in fun and entertaining, mostly small talk conversations, yet with little access to the reasons for successful matches, and therefore little access to reflections on the dialogue.

The interdisciplinary collaborative skills of students, as expected, turned out to be a critical path for success or failure. The same can be anticipated for any intended future implementation of similar material with virtual characters, given that knowledge about basics of AI and state machines as well as educational concepts have to liaise.

4 Conclusions

With regard to the conceptual assumptions exposed in the introduction, the following conclusions can be made for the use of the presented concepts for learning and knowledge media:

The project actually started out with a goal to create virtual characters that can teach and tell stories and facts, which primarily supported a similar conceptual model as of a classic "virtual guide". Coupled with this concept, the educationalists even showed a reluctant attitude due to the concern that virtual characters could possibly serve as a substitute for teachers. There was a significant change of engagement, when the metaphor of a virtual dollhouse was established instead. Naturally, also the complexity of the task was a barrier that needed to be overcome in the first instance. Furthermore, during the design phase, first informal evaluations of characters have been performed with children. A significant amount of girls and boys showed preferences for characters that rather represent a co-learner or a learner, than a teacher.

Another surprising perception was the positive acceptance of keyboard-based text input. Though there are known usability issues such as typing speed and errors, in fact it helps construct knowledge instead of providing only indiscriminate choices, which would be common to most games and e-learning applications with mouse interaction. From the point of view of teaching mathematics in primary school, these are in fact basic requirements, well supported by the functionality of the platform. The system must support an active dispute with the learning matter. Therefore, interactivity (particularly verbal interactivity) of the learners has a high priority. In the math classroom, language is most important. For example, the teacher phrases problems, explains algorithms, and expresses mathematical ideas verbally at first. In the NCTM process standard of "Communication", communication is characterized as "… an essential part of mathematics and mathematics education. It is a way of sharing ideas and clarifying understanding. The communication process also helps build meaning and permanence for ideas and makes them public." [National Council of Teachers of Mathematics 2000].

In summary, a big potential was seen in the integration of small interactive conversations like the ones created within an overall "blended learning" strategy. Mapped onto the interaction concepts outlined in the first chapter, the exploration aspect gained increasing interest, focusing on the medium as an extension of one's imagination while simulating hypotheses. Here, a central learning effect was the active construction of the hypothetical conversations by the students. In addition, training effects occur whilst using language as the interface, testing dialogue moves while typing, providing the construction of practical and tacit knowledge. From a learning point of view, this appeared to be a more promising application of virtual characters to education, than the all-knowing virtual tutor presenting facts.

Acknowledgements

Thanks to Rose Vogel and Wolfgang Müller, for the fruitful interdisciplinary cooperation and discussion on educational aspects; and platform credits to the consortium of the art-E-fact project [Artefact 2005], especially to ZGDV, Darmstadt Germany.

References

ALICE 2005. *Alice Foundation.* Online: http://www.alicebot.org.

ARTEFACT 2005. *art-E-fact project.* Online: http://www.art-e-fact.org

CASSELL, J., BICKMORE, T., BILLINGHURST, M., CAMPBELL, L., CHANG, K., VILHJÁLMSSON, H., YAN, H. 1999. Embodiment in Conversational Interfaces: Rea. In *ACM CHI 99 Conference Proceedings*, Pittsburgh.

FRASCA, G. 2003. Ludologists love stories, too: notes from a debate that never took place. In *Level Up, DIGRA Conference Proceedings*, Utrecht.

ISAGA 2004. *Bridging the Gap: Transforming Knowledge into Action through Gaming and Simulation.* ISAGA annual conference. Online: http://sagsaga.org/isaga2004/

MATEAS, M., STERN, A. 2003. Integrating Plot, Character and Natural Language Processing in the Interactive Drama Façade. In *Proceedings of TIDSE 2003 Conference*, Darmstadt.

MÜLLER, W., SPIERLING, U., WEISS, S. 2005. Synchronizing Natural Language Conversation between Users and Multiple Agents in Interactive Storytelling Applications. In *Proceedings TESI 2005*, Kent, UK

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS 2000. Standards for School Mathematics Communication. *Principles and Standards for School Mathematics,* Chapter 3. Online: http://standards.nctm.org/document/chapter3/comm.htm.

OZ 2002. *Oz Project Publications.* Online: http://www-2.cs.cmu.edu/afs/cs.cmu.edu/project/oz/web/papers.html

SEARLE, J.R. 1975. A taxonomy of illocutionary acts. *Language, Mind and Knowledge*, Minnesota Studies in the Philosophy of Science.

SPIERLING, U., and IURGEL, I. 2003. "Just Talking About Art" - Creating Virtual Storytelling Experiences in Mixed Reality. In *Virtual Storytelling, Proceedings ICVS 2003*, Springer Berlin-Heidelberg.