From Simulated Dialogues to Interactive Performances

Elisabeth André¹, Thomas Rist², and Stephan Baldes²

¹Universität Augsburg, Eichleitnerstr. 30, D-86135 Augsburg, Germany andre@informatik.uni-augsburg.de ²DFKI GmbH, Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany {rist, baldes}@dfki.de

Abstract. In this contribution, we argue in favor of a shift from applications with single presentation agents towards flexible performances given by a team of characters as a new presentation style. We will illustrate our approach by means of two subsequent versions of a test-bed called the "Inhabited Market Place" (IMP1 and IMP2). In IMP1, the attribute "flexible" refers to the system's ability to adapt a presentation to the needs and preferences of a particular user. In IMP2, flexibility additionally refers to the user's option of actively participating in a computer-based performance and influencing the behavior of the involved characters at runtime. While a plan-based approach has proven appropriate in both versions to automatically control the behavior of the agents, IMP2 calls for highly reactive and distributed behavior planning.

1 Introduction

Lifelike characters, or animated agents, provide a promising option for interface development as they allow us to draw on communication and interaction styles with which humans are already familiar. During the last years, animated characters have been used in a number of different application fields including educational software, help systems, and virtual representatives on commercial web pages that act as product presenters and sales assistants (see [1] for an overview). Most of these applications assume settings in which the agent addresses the user directly as if it were a face-to-face conversation between human beings. Such a setting seems quite appropriate for a number of applications that draw on a distinguished agent-user relationship. For example, an agent may serve as a personal guide or assistant in information spaces like the world-wide web.

However, there are also situations in which the emulation of a direct agent-to-user communication - from the perspective of the user - is not necessarily the most effective and most convenient way to present information. For example, an empirical study by Craig and colleagues [5] suggests that indirect interaction can have a positive effect on the user's performance. They found that, in tutoring sessions, users who overheard dialogues between virtual tutors and tutees, subsequently asked significantly more questions and also memorized the information significantly better. In other situations the user may just feel more comfortable in the role of an observer rather than in the role of an active questioner.

In this paper we propose a shift from single character settings towards interactive performances given by a team of characters as a new form of presentation. The use of presentation teams bears a number of advantages. First of all, they enrich the repertoire of possible communication strategies. For example, they allow us to convey certain rhetorical relationships, such as pros and cons, in a more canonical manner. Furthermore, they can serve as a rhetorical device that allows for a reinforcement of beliefs. For instance, they enable us to repeat the same piece of information in a less monotonous and perhaps more convincing manner simply by employing different agents to convey it. Last but not least, the single members of a presentation team can serve as indices which help the user to organize the conveyed information. For instance, we may convey meta-information, such as the origin of information, or present information from different points of view, e.g. from the point of view of a businessman or the point of view of a traveler.

Our proposal is inspired by the evolution of TV commercials over the past 40 years. A typical commercial of the early days featured a sales person who presented a product by enumerating its positive features – quite similar to what synthetic characters do on web pages today. On TV, however, this format has almost completely been replaced by formats that draw on the concept of short, but entertaining episodes or sketches. Typically, such performances embed product information into a narrative context that involves two or more human actors. One of the reasons that may have contributed to the evolution of commercial formats is certainly the fact that episodic formats offer a much richer basis compared to the plain enumeration of product features, and thus meet the commercial industry's high demand for originality and unseen spots. In this context, we also refer to psychological work that emphasizes the strength of episodic memory [21].

The goal of our research, however, is not to simply imitate episodic formats from TV commercials for information presentation. Rather, we aim at new formats which bring in *adaptivity* and *interactivity*. By adaptivity, we mean a system's ability to adapt a presentation to the needs and preferences of a particular user. Interactivity refers to the user's option of actively participating in a performance. We will illustrate our approach by means of two subsequent versions of a test-bed called the "Inhabited Market Place" (IMP). We will use the abbreviations IMP1 and IMP2 to refer to version 1 or 2 respectively.

2 The Inhabited Market Place

As the name indicates, the Inhabited Market Place [2] is a virtual place in which seller agents provide product information to potential buyer agents in form of a typical multi-party sales dialogue. Figure 1 illustrates the scenario: the characters Merlin and Robby on the left-hand side play the role of car dealers who present information about the displayed car, and answer questions posed by the two buyer agents Peedy and James on the right-hand side. The user has the option of joining the discussion as well – either as a buyer or a seller. In Fig. 1, the user is represented by the agent in the middle. 1



Fig. 1. The Inhabited Market Place

2.1 Structure and Representation of the Product Data

Part of the domain knowledge is an ordinary product database, e.g., organized in the form of an n-dimensional attribute vector per product. In our current scenario, the products are cars with attributes, such as model type, maximum speed, horsepower, fuel consumption, price, air conditioning, electric window lifters, airbag type etc. Thus, to a large extent, the contents of the database determines what an agent can say about a product. However, products and their attributes are described in a technical language which the user may not be familiar with. Therefore, it seems much more

¹ Apart from the user's avatar which was created by DFKI GmbH, all other agents have been taken from the Microsoft Agent Ring (see http://www.msagentring.org).

appropriate to maintain a further description of the products - one that reflects the impact of the product attributes on the value dimensions of potential customers. Such an approach can be modeled in the framework of multi-attribute utility theory (e.g. see [24]), and has already been used for the identification of customer profiles in an electronic bourse for used cars [16]. In this project, the car database was provided from a large German/American car producer and retailer, whereas the value dimensions for the product "car" have been adopted from a study of the German car market [19] that suggests that safety, economy, comfort, sportiness, prestige, family and environmental friendliness are the most relevant. In addition, it was represented how difficult it is to infer such implications. The work presented here follows this approach even though we employ a simplified model. For instance, we use the expressions:

```
FACT value "consumptioncar1" 8;
FACT polarity "consumptioncar1" "environment" "neg";
FACT difficulty "consumptioncar1" "environment" "low";
```

to represent that a certain car consumes 8 liters, that this fact has a negative impact on the dimension "environment" and that this implication is not difficult to infer.

2.2 Setting the Parameters for the Presentation

In the scenario shown in Fig. 1, the salesman Merlin is trying to convince the buyers of the potential benefits of the displayed car. From the point of view of the system, the presentation goal is to provide the user with facts about a certain car. However, the presentation is neither just a mere enumeration of the plain facts about the car, nor does it assume a fixed course of the dialogue between the involved agents. Rather, IMP1 supports the concept of adaptivity. It allows the user to specify prior to a presentation (a) the agents' role, (b) their attitude towards the product (in our case a car), (c) their initial status, (d) their personality profile and (e) their interests. Taking into account these settings, a variety of different sales dialogues will be generated for one and the same product. Figure 2 shows the interface that allows the user to input these settings.

The interest profile is used to determine those attributes of the car that should be addressed in the dialogue. Depending on the agents' attitude towards the product, their status and personality, positive or negative evaluations about these attributes are added.

To model the characters' personality, we adopt the so-called Five-Factor Model [15]. The FFM is a descriptive model, with the five dimensions (*Extraversion, Agree-ableness, Conscientiousness, Neuroticism*, and *Openness*) being derived from a factor analysis of a large number of self- and peer reports on personality-relevant adjectives. We decided to focus on the dimensions: *Extraversion, Agreeableness*, and *Neuroticism* which seem to be most relevant for social interactions.

3 Automated Script Generation for Simulated Sales Dialogues

In IMP I, the system takes on the role of a screen writer that scripts the behavior of a group of actors that participate in a dialogue. We follow a communication-theoretic view and consider the automated generation of such scripts a planning task. The implementation of the planning approach is based on the JavaTM-based JAM agent architecture [9]. To model the knowledge used for script generation, we defined planoperators that code a decomposition of a complex communicative goal into dialogue acts for the single agents.



Fig. 2. Interface for Specifying a Character Profile

An example of a dialogue operator is listed in Fig. 3. It represents a scenario where two agents discuss a feature of an object. The operator only applies if the feature has a negative impact on any value dimension and if this relationship can be easily inferred. According to the operator, any disagreeable buyer produces a negative comment referring to this dimension (NegativeResponse). The negative response is followed by a response from a seller (ResponseNegativeResp).

One possible response is listed in Fig. 4. It only applies if there is an attribute that has a positive impact on the dimension under discussion. In this case, the seller first tells the buyer(s) that it disagrees and then lists attributes with a positive impact on the dimension. Note that our plan operators include both the propositional contents of an utterance and its communicative function. This is in line with Cassell and colleagues, who regard conversational behaviors as fulfilling propositional and interactional conversational functions [4]. For instance, we explicitly represent that "Bad for the " \$dimension "?" is a response to a negative comment.

NAME: "DiscussValue1" GOAL: PERFORM DiscussValue \$attribute; PRECONDITION: FACT polarity \$attribute \$dimension "neg"; FACT difficulty \$attribute \$dimension "low"; FACT Buyer \$buyer; FACT Disagreeable \$buyer; FACT Seller \$seller; BODY: PERFORM NegativeResponse \$buyer \$dimension; PERFORM ResponseNegativeResp \$seller \$attribute \$dimension;

Fig. 3. Example of a dialogue operator for discussing an attribute value

Fig. 4. Example of a plan operator for responding to a negative comment

The character's profile is considered by treating it as an additional filter during the selection instantiation and rendering of dialogue strategies. In particular, we define specific dialogue strategies for characters of a certain personality and formulate constraints that restrict their applicability. The script planning mechanism ensures that the generated scripts entail for each character only role-appropriate or role-neutral behaviors that do not conflict with its personality profile. For instance, a customer in a sales situation usually tries to get information on a certain product in order to make a decision, while the seller aims at presenting this product in a positive light. In contrast to an extrovert agent, an introvert agent will less likely take the initiative in a dialogue.

To illustrate the adaptive features of our system, let's have a look at the two sample dialogues listed in Fig. 5. For expository reasons, we use extreme parameter settings so that differences in the behavior of the characters are readily distinguishable. The dialogues partially discuss the same car attributes, but from different points of view. In both cases, one of the buyers criticizes the high gas consumption of the car. But in the first case, it is concerned about the environment, while, in the second case, it is thinking of the high costs.

Robby: Role: seller; Personality: agreeable, extravert; Interests: environmental issues Peedy: Role: customer; Personality: disagreeable, introvert; Interests: environmental issues

Peedy:

How much gas does it consume? ;;; wants to know more about gas ;;; consumption because this feature has an ::: impact on "environment," which is ::; important to him **Robby:** It consumes 8 liters per 100 km. ;;; retrieves the value from the car database Peedv: Isn't that bad for the environment? ;;; the value of the attribute "gas ;;; consumption" has a negative impact on ;;; "environment," Peedy is disagreeable ;;; and therefore makes a negative ;;; comment, less direct speech since it is ::: introvert **Robby:** Bad for the environment? ::: questions negative impact It has a catalytic converter. It is made of recyclable material. ;;; provides counterarguments

•••

Robby:

Role: seller; Personality: agreeable, extravert; Interests: sportiness Peedy: Role: customer; Personality: disagreeable, extrovert; Interests: economy

Peedy:

How much gas does it consume? ;;; wants to know more about gas ;;; consumption because this feature has an ::: impact on the dimension "economy," ;;; which is important to him **Robby:** It consumes 8 liters per 100 km. ;;; retrieves the value from the car database Peedv: I'm worrying about the running costs. ;;; the value of the attribute "gas ;;; consumption" has a negative impact on ;;; "economy," Peedy is disagreeable and ;;; therefore makes a negative comment **Robby:** Forget the running costs. Think of the prestige. ;;; tries to distract the buyer by ;;; mentioning any positive implication. ;;; "prestige" has been chosen since there ;;; is no other easier-to-infer positive ;;; implication

Fig. 5. Dialogues for different parameter settings

4 Structuring Interactive Performances

The presentation task and scenario of IMP II are similar to IMP I. In addition, IMP II provides the user with the option of taking an active role in the performance if she or he wishes to do so. If not, however, the characters will give a performance on their own – maybe encouraging the user to give feedback from time to time. At each point in time, the user has the option of joining the discussion again. The novelty of the approach lies in the fact that it allows the user to dynamically switch between active and passive viewing styles. Such a scenario bears a lot of similarities to improv-

isational theatre (cf. [10]). First of all, there is no pre-defined script. Instead the dialogue between the user and the characters evolves while time progresses. Furthermore, the scenario is open-ended. Neither the characters nor the users are able to tell what exactly may happen next.

Since user reactions cannot be anticipated prior to a presentation, it is no longer possible to pre-script utterances. Instead scripting has to be done at runtime, e.g. either by a centralized script writing component or by the single agents themselves. For IMP II, we decided to use a self-scripting approach and realize our characters as autonomous agents. That is, the behavior of each agent is triggered by events occurring in the scene and the dialogue contributions of the other agents and the user. Fig. 6 provides an overview of the architecture we developed for IMP II.

The *Jimpro* module implements an interface between the agents' body (which is realized by the Microsoft Agent Technology [17]) and their mind (which is realized by different JAM clients). It consists of an agent server which registers all agents that participate in a conversation and handles the communication between them. The agent handler is responsible for the execution of the elementary animations and speaking actions provided by the underlying audio-visual interface (in our case the Microsoft agents). In addition, it informs the JAM clients about the status of the executed actions and the audio-visual interface.

The *Dialogue Management Component* maintains the dialogue history and a list of dialogue goals stored on a goal board that still need to be addressed by the agents. All agents have access to the goal board and may apply for the right to accomplish goals. Which agent will succeed depends on the given dialogue protocol. As a first step, we implemented a protocol that supports well-organized conversations. Among other things, this protocol prescribes that questions should be addressed before new dialogue contributions are made (unless none of the agents is able to provide an answer). Furthermore, if an agent is addressed directly, he will get the right of speaking with a higher probability than any other agent. In addition, we consider the agents' status and personality when allocating dialogue turns. For instance, extrovert agents with a high status get a word in with a higher probability than introvert agents with a low status.

The *JAM clients* represent the agents' mind. We assign each agent its own reactive planner and a data base which contains its world knowledge. Furthermore, each agent has a repertoire of dialogue strategies at its disposal that are coded as plan operators. We were able to reuse most dialogue operators from IMP I that represent the behavior of a single agent (like the plan operator listed in Fig. 4). However, these operators are no longer employed by a central script-writing component, but now belong to an agent's individual repertoire of dialogue behaviors. Plan operators that specify a dialogue sequence between several agents, such as the plan operator listed in Fig. 3, have been replaced in IMP II by rules that map dialogue events onto dialogue goals for the single agents. For instance, if one of the agents (or the human user) asks a question, the dialogue management component puts the goal to respond to that question onto the goal board.

Dialogue contributions result from autonomous characters trying to achieve their individual goals. The goals of the single agents are derived from their role and personality profile. For instance, an agent that takes on the role of a car seller is initialized with the goal to sell a car. In addition, we associate with agreeable and



Fig. 6. Architecture for IMP II

extrovert agents the goal to perform polite behaviors which means among other things that they start a greeting behavior whenever they come across another agent for the first time. Unless an agent is both disagreeable and introvert, it adopts the goal to respond to a question whenever it kknows the answer.

As shown in Fig. 6, the user's avatar is assigned a JAM client as well. This relieves the user from the burden to specify the behavior of his or her avatar to the last detail. The JAM client is triggered by the natural-language utterances of the user which are analyzed by Spin, a Java-based template-matching tool developed within the SmartKom project [22].

5 Related Work

The Agneta and Frida system [8] incorporates narratives into a Web environment by placing two characters on the user's desktop. These characters watch the user during

the browsing process and make comments on the visited Web pages. Unlike our approach, the system relies on pre-authored scripts, and no generative mechanism is employed. Consequently, the system operates on predefined Web pages only.

An earlier system by Cassell and colleagues automatically generates and animates dialogues between a bank teller and a bank employee with appropriate synchronized speech, intonation, facial expressions, and hand gestures [3]. However, their focus is on the communicative function of an utterance and not on the personality and the emotions of the single speakers. Furthermore, they do not aim to convey information from different points of view but restrict themselves to a question-answering dialogue between the two animated agents.

Walker and colleagues [23] concentrate on the linguistic capabilities of computer characters and examine how social factors influence the semantic content, the syntactic form and the acoustic realization of conversations. The generation of their dialogues is essentially influenced by the power the listener has on the speaker and the social distance between them. This approach has been later extended by Prendinger and colleagues to create animated dialogues for the Microsoft agents [18].

Our work was heavily inspired by research on interactive drama that aims at integrating a user in a scenario – either as an audience member or an active participant. To allow for user interaction, systems usually incorporate decision points in a narrative-style script [14] or model their characters as autonomous agents that select and instantiate actions under consideration of dramatic constraints, such as the plot of a story or the characters' role and personality [7]. The integration of dramaturgical elements has been proven useful for a large variety of applications. In particular, developers of educational software are increasingly employing narrative concepts to structure their pedagogical material. The spectrum ranges from virtual puppet theatres [11,12] for children to educational soap [13] and team training [20].

6 Conclusion

In this paper, we proposed performances given by a team of animated agents as a new presentation style. Infotainment and edutainment transmissions on TV as well as advertisement clips are examples that demonstrate how information can be conveyed in an appealing manner by multiple presenters with complementary characters and role castings. However, our approach distinguishes from conventional TV presentations by at least two features: *adaptivity* and *interactivity*. We demonstrated this by means of two subsequent versions of a test bed called the Inhabited Market Place.

In IMP I, we automatically generated scripts for sales dialogues depending on a number of character-related parameters to be specified prior to a presentation. The scripting approach bears the advantage that it enables the generation of coherent dialogues. It requires, however, that all the knowledge to be communicated to the audience is known in advance. Consequently, it is less suitable in situations where the agents have to immediately respond to events at presentation runtime, such as new incoming information to be presented or user interactions.

In IMP II, we moved to a character-centered approach. Instead of specifying the agents' behavior to the last detail, we just provide a character with a description of its role and profile according to which it has to behave at presentation runtime. Such an

approach seems appropriate for scenarios that require immediate responses to external events, such as user interactions. It is, however, more difficult to ensure the coherence of a dialogue since no global organization of the information is possible.

In the future, we will make more extensive use of dramaturgical elements in order to achieve more interesting interactions. Inspired by our work on the Puppet project [11], we are currently investigating a dramaturgy framework which goes back to Greimer's ACTANT model [6]. A major element of the framework is that of an underlying conflict which is established by introducing a protagonist, that persecutes a certain goal, and a second character, the antagonist, that tries to accomplish a counter goal. Both the protagonist and the antagonist may be supported by one or more helpers. Once started, a certain "dramatic story" would unfold over time just by having the involved actors play their assigned roles. In the case of the Inhabited Market Place, we might model a buyer and a customer with conflicting interests. While the seller tries to present a car in a positive light, the customer persecutes the opposite goal – namely to point out the weaknesses of the car. In addition, we foresee a helper character who is to support the virtual customer and is played by the user. As a helper agent, the user may interact in the following ways: He or she may support the virtual customer directly by confirming its statements or expressing approval. In response to a seller's statement, he or she may utter disbelief or mention understanding problems. The Inhabited Market Place II can be seen as a test bed which allows users to experiment with different dramaturgical elements in order to find out what kinds of setting tend to result into interesting improvisations.

Acknowledgements. The work described here has been partially funded by the BMBF project MIAU and the EU projects MagiCster and NECA. We are grateful to Peter Rist for the graphical design of Mr. Smith. We would also like to thank Bastian Blankenberg and Martin Klesen for the implementation of Jimpro and the interface for IMP I.

References

- 1. André E. (1999). Applied Artificial Intelligence Journal, Special Double Issue on Animated Interface Agents, Vol. 13, No. 4-5.
- André, E., Rist, T., van Mulken, S., Klesen, M. and Baldes, S. (2000). The Automated Design of Believable Dialogues for Animated Presentation Teams. In: Cassell, J., Sullivan, J., Prevost, S. and Churchill, E. (eds.): *Embodied Conversational Agents*, 220-255, Cambridge, MA: MIT Press.
- Cassell, J., C. Pelachaud, N. I. Badler, M. Steedman, B. Achorn, T. Becket, B. Douville, S. Prevost, and M. Stone. (1994). Animated conversation: Rule-based generation of facial expression, gesture and spoken intonation for multiple conversational agents. *Computer Graphics* (SIGGRAPH '94 Proceedings), 28(4):413-420.
- 4. Cassell, J., Bickmore, T., Camphell, L., Vilhjalmsson, H. and Yan, H. (2000). The human conversation as a system framework: Designing embodied conversational agents. In: Cassell, J., Sullivan, J., Prevost, S. and Churchill, E. (eds.): *Embodied Conversational Agents*, 29-63, Cambridge, MA: MIT Press.

- Craig, S. D, B. Gholson, M. H. Garzon, X. Hu, W. Marks, P. Wiemer-Hastings, and Z. Lu. 1999. Auto Tutor and Otto Tudor. In *AIED-Workshop on Animated and Personified Pedagogical Agents*, 25–30. Le Mans, France.
- 6. Greimers, A. and Courtes, J. (1982). *Semiotics and Language: An Analytical Dictionary*. Bloomington, IN: Indiana University Press.
- 7. Hayes-Roth, B., van Gent, R. and Huber, D. (1997). Acting in character. In: Trappl, R. and Petta, P. (eds.): *Creating personalities for synthetic actors*, 92–112. New York: Springer.
- Höök, K., Persson, P. and Sjölinder, M. (2000). Evaluating Users' Experience of a Character-enhanced Information Space. Artificial Intelligence Communications, 13(3), 195-212.
- 9. Huber, M. (1999). JAM: A BDI-theoretic mobile agent architecture. Proc. of the Third Conference on *Autonomous Agents*, 236–243. New York: ACM Press.
- 10. Johnstone, K. (1989). IMPRO: Improvisation and the Theatre. Routledge: New York.
- Klesen, M., Szatkowski, J. and Lehmann, N. (2001). A Dramatised Actant Model for Interactive Improvisational Plays. In: Proceedings of the Third International Workshop on Intelligent Virtual Agents, 181-194, New York: Springer.
- 12. Machado, I., Paiva, A. and Prada, R. (2001). Is the Wolf Angry or ... Just Hungry? In: Proc. of the Fifth Conference on *Autonomous Agents*, 370–376. New York: ACM Press.
- 13. Marsella, S.C., Johnson, W.L. and LaBore, C. (2000). Interactive Pedagogical Drama. In: Proc. of the Fourth Conference on *Autonomous Agents*, 301–308. New York: ACM Press.
- M. Mateas, M. (1997). An Oz-Centric Review of Interactive Drama and Believable Agents. Technical Report CMU-CS-97-156, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.
- 15. McCrae, R. R., and John, O. P. (1992). An introduction to the five-factor model and its applications. Special Issue: The five-factor model: Issues and applications. Journal of Personality 60: 175-215, 1992.
- Mehlmann, O., Landvogt, L., Jameson, A., Rist, T. and Schäfer, R. (1998). Einsatz Bayes'scher Netze zur Identifikation von Kundenwünschen im Internet. *Künstliche Intelligenz* 3(98):43–48.
- 17. Microsoft Agent: Software Development Kit (1999). Microsoft Press, Redmond Washington.
- 18. Prendinger, H. and Ishizuka, M. (2001). Social Role Awareness in Animated Agents. In: Proc. of the Fifth Conference on *Autonomous Agents*, 270–377. New York: ACM Press.
- 19. Spiegel-Verlag. (1993). SPIEGEL-Dokumentation: Auto, Verkehr und Umwelt. Hamburg: Augstein.
- Swartout, W., Hill, R., Gratch, J., Johnson, W.L., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J., Rickel, J., Thiébaux, M., Tuch, L., Whitney, R. and Douglas, J. (2001). Toward the Holodeck: Integrating Graphics, Sound, Character and Story. In: Proc. of the Fifth Conference on *Autonomous Agents*, 409–416. New York: ACM Press.
- 21. Tulving, E. (1983). Elements of Episodic Memory. Oxford: Clarendon Press.
- 22. Wahlster, W., Reithinger, N. and Blocher, A. (2001). SmartKom: Multimodal Kommunikation with a Life-Like Character. In: Proceedings of *Eurospeech 2001*, Vol. 3, 1547-1550.
- Walker, M., Cahn, J. and S. J. Whittaker, S.J. (1997). Improving linguistic style: Social and affective bases for agent personality. In *Proceedings of Autonomous Agents*'97, 96– 105. Marina del Ray, Calif.: ACM Press.
- 24. von Winterfeldt, D., and W. Edwards. (1986). Decision analysis and behavioral research. Cambridge: Cambridge University Press.