The Future of Interactive Drama
Nicolas Szilas
Macquarie University
Department of Computing
Macquarie University NSW 2109
+61 2 9850 9571
nicolas@ics.mq.edu.au

ABSTRACT
Interactive Drama is the ultimate challenge of digital entertainment. In this paper, from our seven year experience in Interactive Drama, we try to shape the history of the field and envision what will be (or should be) the future of this history. Two main directions in particular are stressed, because we feel that the success of Interactive Drama lies in these two directions. The first one concerns the architecture of systems and how it would manage both narrative constraints and character's intelligence, believability and roundness. The second one focuses on project management by sketching a methodology of co-design for Interactive Drama.

Categories and Subject Descriptors

General Terms

Keywords
Human Computer Interaction, Narrative Intelligence, Interactive Narrative, Interactive Drama, Narrative Structures.

1. INTRODUCTION
Interactive Drama (ID) is a topic that is easier to explain to an average video gamer than to an academic researcher... Let us start with the gamer definition:
Interactive Drama is a solo adventure game where you really influence the story.

Then, the academic definition would be:
Interactive Drama is a narrative genre on computer where the user is one main character in the story and the other characters and events are automated through a program written by an author. Being a character implies choosing all narrative actions for this character.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Conference '04, Month 1–2, 2004, City, State, Country.
Copyright 2004 ACM 1-58113-000-0/00/0004…$5.00.

In this definition and as in the rest of the paper the term "narrative" does not refer to a simple "recounting of a sequence of events" [14] but to a recounting which contains features that turns it into what is called a "story" in the basic sense of the word (see[1] and [6] for detailed definitions). The narrative actions in the above definition are actions which have a significant impact on the story. In Interactive Drama, non narrative actions such as walking into a room do not need to be controlled by the user.

Interactive Drama, as defined above, does not yet exist. Research prototypes have been developed [8][11][20][24][28] but they do not yet exhibit convincing artistic or entertaining value. The only released prototype, Façade [12], does not meet the criteria above, because the characters have no deep understanding of user's utterances (e.g. the user is not able to transmit a specific piece of knowledge to other characters). The question is: when any of these prototypes will constitute an entertaining and playable interactive drama, in the sense defined above?

All these innovative projects on ID span over many years and it is becoming clear that building an ID from scratch is not an achievable goal within the three year duration of a PhD. To make it possible to create ID within a shorter period of time, we believe that it is necessary to overcome two types of obstacles:

- Technical problem: many projects focus on building a suitable technical environment, which includes graphics, sound, interaction, characters, etc. but then lack resource on the algorithmic issues related to narrative and interactivity.

- Conceptual problem: ID is neither a pure technical problem nor a pure creative problem. It combines these two sets of approaches in a radically new way that is difficult to grasp, especially at the beginning of a project.

The first issue requires the reuse of algorithms and code between different projects. We believe that this will be the case in the coming years, because more and more research institutions promote open source outcomes (see [5] for example). The current use of game engines [2][7][16][28] also makes it possible to reach good graphical quality with less effort.

This paper deals with the second issue, the conceptual one. From our experience in building ID, we would like to provide a picture of the main conceptual trends in ID and anticipate the new avenues worth exploring in order to achieve ID in a reasonable number of years.

We have divided this picture into pseudo time slices (Section 2): the past, the present and the near future, the future being described in sections 3 and 4. In this categorization if a system belongs to the “past” or “present” category, it does not mean it is
older, neither than the system from the near future are less developed. The belonging to one category or another means that a system is based on some principles that appeared at some period of time.

This original way of categorizing systems has been inspired by art history. Indeed, in the art domain, artworks created at the same period belong to completely different artistic currents. In science as well, to properly understand a field, one needs to understand the different but coexisting streams of ideas involved in that field. In Interactive Drama, some systems are based on "past ideas", more anchored in "old media" while other, we believe, suggest more novel views of Interactive Drama, in term of story management.

Sections 3 and 4 are almost independent. They describe two major research avenues which are overlooked in current systems. They could be used a starting point of a new generation of Interactive Drama systems.

2. HISTORY OF THE FIELD

2.1 The Past

The most straightforward way to think of ID is in terms of branching. In branching narrative, a story is seen as a hypermedia. It is composed of (and authored as) a set of fragments interconnected by links. The user’s action consists of navigating from fragments to fragments, following the links. In particular, the hypermedia can be designed so as the user plays the role of one character. The navigation between fragments correspond to choices made by the character. Choose your own adventure books are typical branching narratives of this kind. Other genres, while allowing more interaction to the user are implicitly based on branching, as far as the story is concerned. It is the case of some adventure video games having several endings (e.g. Blade Runner for PC).

Limits of branching are well known. The more interactivity it is given to the user, the more effort is required from the author who has to design all branches. Thus, building an ID with branching makes the number of fragments to write intractable.

Note that there exists a variation of branching, conditional or adaptive branching. In that case, the existence of a link between two fragments is calculated during the execution time, according to author-defined laws. This allows a much better control of the user experience within a given structure, but does not solve the problem of ID because the fragments themselves still have to be manually written.

2.2 The Present

Several current systems are based on the notion of variable fragment: Each fragment is defined in an abstract way rather than being completely scripted.

For example, Brian Magerko proposes to define a story in terms of a linear sequence of generic scenes: a scene is defined with variables, which are instantiated during execution [11]. The same principle of genericity has been used back in the seventies [10], for generating new tales from the Propp model [15].

Facade [12][22] is another example of abstract fragments, where the scenes (called beats) are predefined but variable in the sense that they can be interrupted and interwoven so that the flow of events reacts to user's actions.

These systems are also characterized by the fact that the notion of link between fragments has been replaced by conditions of triggering, what is called temporal variability in [11].

This approach suffers from two main issues:

- Interactivity is limited because the grain of each fragment is large, so the user only retrieves a prewritten scene, rather than creating his/her own situation. According to Marie-Laure Ryan, the user is more an "active observer rather that being cast as the main protagonist" [19].

- To obtain more interactivity one has to write a lot of scenes, which is practically quite problematic.

2.3 The Near Future

The next step consists in using temporal variability and content variability at a smaller scale, the level of actions. Agent based systems are used to simulate characters, from low level abilities (animation, reactive behaviors, emotions) to higher level abilities (strategic planning, dialog) [3][7]. However, these systems lack narrative guidance, because a narrative is not exactly a simulation of characters, as we have argued elsewhere [23], because a narrative is structured as a global message towards it audience.

Other systems implement a drama management feature, in order to dynamically generate a plot, according to the user's actions. One of the first and most advanced system of this kind is Defacto, where actions are calculated from rules implementing narrative strategies such as: “Accept as plot developments only character intervention affiliated with the current storyline goal (unity of plot)” [20].

The system developed in the Liquid Narrative Group also manages a narrative at the level of action by using a planning algorithm able to guide a story by anticipating all possible outcomes of a story and choosing the best one, according to various constraints including authoring constraints [17][28][29].

Our own system called IDtension also intends to guide a narrative according to narrative criteria. These criteria consist of “effects” that the system wishes to produce towards the audience. The system is also based on generic narrative types of actions such as “inform”, “encourage”, etc. which allows greater interactivity and less authoring work [23][24][25]. Typically, by only writing a few story elements such as tasks (steal, offer), characters (John, Mary, Bob), and objects (key, ring), the system can generate dozens of different complex actions, such as "John dissuades Mary to steal the key from Bob".

These systems however suffer from a poor narrative quality compared to fragment-based approaches. This might be solved by working extensively on more complete agents and richer narrative criteria for guidance. But it is our opinion that some conceptual changes will be needed in the future, in order to accelerate the development of interesting, highly interactive drama.

3. DRAMA MANAGEMENT AND AGENTS

From the observation of existing systems for interactive narrative and drama, the relation between the narrative management and the character's behavior has always been very simplistic, and we conclude that this might be a fundamental limitation of current systems.

In particular, the relation between the “drama manager” (DM) and the “intelligent agents” (IA) is always hierarchical (the DM sends commands to the IAs) and the variations between systems lies in
the level of the communication. In the Oz Project, an early system for ID [3], the DM intervenes punctually by assigning goals to IAs. In that case, the communication between the DM and the IAs happens at the higher levels of goals. Other systems like Façade or IDtension have considered that the DM should intervene at a lower level, while the intelligence of the agents should be limited to reactive and emotional behaviors [12][23], similarly to other systems like Defacto [20] or Mimesis [28][29]. But giving more room to the DM has created a new set of problems:

- In action-based systems (Mimesis [29], Defacto [16], IDtension [24]), the characters only handle strategic goals, without being able to react to their environment as reactive agents
- In scene-based systems (Façade [12]), the agents do not reason at all, which limits the interaction (e.g. it is impossible to give a specific piece of information to a character).

These problems occur because the relation between the DM and the IAs remains hierarchical as depicted in Figure 1.

Figure 1.: Hierarchical view of drama management

The solution consists in rethinking the relationship between the DM and IAs in terms of cooperation rather than subordination. It is obvious that both the IAs and the DM, that is both character reasoning and narrative reasoning, are concerned with high level beliefs, actions and goals. Both also should manage characters’ emotions. The problem mentioned above comes from the attempt to a priori distribute some control between the two entities, in a hierarchical manner, while this control should be negotiated.

What kind of negotiation should take place between these entities? Two schemes of communication are suggested.

In the first scheme, the initiative comes from the IA, as depicted in Fig. 2. The communication cycle is as follows:

1. Each agent calculates a set of actions that could rationally be performed given the current fictional environment. It gives a score of believability for each action. Before performing any of these actions, they send them to the DM, with the corresponding motivation for each action. Motivations vary according to the agents’ knowledge representation. Typically, it consists of the goals that the action tends to reach or impede, the side effects of this action, the risks related to this action, etc.

2. The DM evaluates the actions based on narrative criteria other than rationality or believability of characters. This is highly dependent of each system. A commonly used criterion is the conflict related to the current action [17][20][23], but several other relevant criteria have been proposed as well such as Suspense, Thought Flow or Unity of Plot [20][24][27][28]. Any kind of algorithm can be used for the evaluation, including algorithms which plan ahead possible actions [27][28]. According to this evaluation, the DM either chooses the most rational action, according to IAs, or select a less rational one but relevant from a narrative point of view. This action is then sent back to the corresponding IA, with a direction. A direction is an explanation of why the action is to be played from a narrative point of view. Typically, it contains the criterion that make the DM select this action.

3. The IA that receives the action plays it, according to the direction. In that case, the IA is not a character but an actor, analogous to a theater actor: It not only acts to fit a character but he is also aware of the fact that the act is part of a narrative and its effect to the audience.

Another scheme consists in giving the initiative to the DM (Fig. 3):

1. The DM calculates a set of actions which are relevant according to narrative criteria. Each action is given a score, quantifying its “narrative interest”. Each action is sent to the corresponding IA.

2. Each IA which receives an action from the DM examines its credibility given its beliefs and goals. At this level, the agent has the possibility to come up with some new facts to make the action credible. For example if it is desirable that the character does not answer the question that the user is asking, then the irruption of a third character in the dialog, with a specific goal, would be suggested by the IA as a new fact. If the action is accepted, it is sent back to the DM along with a score of credibility and a possible list of new facts that should be added to make the action credible.

3. The DM then selects, from all the accepted actions received from the IAs, the most interesting one (see step 1), granted
that the added facts are compatible with the current state of
the world and the narrative criteria. It then sends this action to
the concerned IA with clear directions (see Step 2 of previous
scheme). It also adds to the storyworld the new facts proposed
by the AI, if any. Note that a more sophisticated evaluation
function can be used, to select the action best balanced
between credibility and narrative interest.

4. The IA plays the action like an actor based on given directions
by the DM (see Step 3 of previous scheme).

In the second scheme, it is assumed that the IA will find at least
one credible action in the list of actions suggested by the DM.
Otherwise, a n-step negotiation process has to be set up.

Note that the two schemes could be combined into a single
system.

The architectures corresponding to these two schemes suppose a
different kind of IAs, which are virtual actors rather than virtual
characters. Often, the two terms have been used interchangeably
and the virtual actors have no acting intelligence. One exception
is the ACONF architecture [18], where the IA can choose to play
a sub-part of a narrative plan, according to its own features.

These two schemes which are derived from architectural
considerations are related to existing narrative theories and
practices. Our intention is not to justify this by narrative theories,
because narrative theories are so diverse that, to some extent, any
architecture could be justified by finding the appropriate narrative
theory [24]. Our goal is rather to illustrate the idea of negotiation
between the IAs and the DM with relevant narrative theories and
practices.

First, the question of the relation between plot and characters has
been quite intelligently discussed in a paper from Gérard Genette
entitled “Vraisemblance et Motivation” (plausibility and
motivation) [9]. In this bright text, Genette explains that
fundamentally, a narrative is arbitrary, which means that it follows
its own rules. Then this arbitrary nature is dissimulated by some
causal links, which Genette qualifies as a posteriori justifications.
In other words, the motivations of the characters (why they
behave as they behave) is a justification of narrative-based
actions. This is quite close to our second scheme (Fig. 3) where
the DM proposes some actions and the IAs justify them.

Second we find resonances between the proposed architecture and
the way theatrical improvisation works. Among the practical rules
used by improvisation actors, one is “using contrasts”. It means
that it is worth having two characters on the stage who behave
completely differently in terms of attitude, emotions, social class,
etc. (see [4] for a list of such situations). As a consequence, when
the first actor enters the stage with a certain emotion (e.g. crying),
then the other actor would typically enter the stage with the
opposite emotion (e.g. laughing). Given the quick pace of the
improvisation, this second actor enters the stage without having
taken the time to think of a reason why s/he would have such an
emotion. It is only after a certain amount of time on the
stage, based on the interaction with the first actor, that the second
actor finally finds a justification for his/her behavior and then acts
accordingly. Thus, contrary to a pure character-based approach,
improvisational actors do use narrative constraints. We have here
a working example of the scheme described in Fig. 3. In that case,
the two components are located in a single acting entity.

4. THE AUTHOR IN THE LOOP

The problem of interactive stories was identified more than two
decades ago. It is now a known fact that it can not be solved
solely at the level of writing and artistic design. It requires the use
of Artificial Intelligence to produce narrative events in a
generative manner. This constitutes an ambitious agenda,
covering various fields such as procedural animation, behavioral
modeling, emotional modeling, Natural Language Processing,
computational narrative, Human Computer Interaction. This could
be the reason why Interactive Stories are yet to come. However,
even minimal interactive stories, with basic graphics, texts and
complexity have not been produced so far. Thus, we conjecture
that part of the explanation lies elsewhere, namely in the
methodology to produce the interactive stories.

Building one Interactive Drama requires two types of skills which
can be subcategorized into two more specific skills (Fig. 4):

![Figure 3. Dialogical view of drama management: Drama Manager initiative](image-url)

We believe that the second scheme is the most promising, for the
following reasons:

- The intelligence of the agent is not confined to character's
  rationality but it extends to the agent's ability to justify itself.
  This provides for a good characterization of believability: a
  character is believable when its actions can be explained
  somehow.

- It includes the possibility to add new facts to the storyworld in
  order to justify an action which increases the flexibility of
  narrative in Interactive Drama.

- Even if the initiative is given to the DM, the IAs benefit from
  a high level of reasoning intelligence. But this intelligence is
  used exactly where it is needed to justify the course of
  narrative events, rather than to exhibit some intelligence for
  the sake of intelligence\(^1\).

\(^1\)Intelligence for the sake of intelligence can have a spectacular
value but it is not worth putting efforts in characters' intelligence
just for spectacle.
An Interactive Drama is composed of three components: related to such a methodology for Interactive Drama. However, that understanding is not sufficient in order to build narrative theories is important to the algorithmic designer.

Still, artists need to think in algorithmic terms, what is referred as procedural literacy [13] but Interactive Drama can not rely solely on their ability program entire works. The rest of the paper concerns the goal of creating such tools. Not surprisingly, Façade, the only successful attempt to produce an Interactive Drama beyond the “research demo” phase, has been made almost exclusively by two people who consider themselves as artists, researchers, programmers and designers. This corroborates our statement that artistic skills and technical skills have to be combined at the early stages of design. However, finding a single person who possesses all these combined can be quite difficult because (1) such persons are hard to find, (2) the workload is huge and (3) it prevents valuable authors, in particular great storytellers, to contribute. Thus, we believe that the future of Interactive Drama lies in tools designed for artists but not in their ability program entire works. The rest of the paper concerns the goal of creating such tools.

Still, artists need to think in algorithmic terms, what is referred as procedural literacy [13] but Interactive Drama can not rely solely on procedural literacy. Symmetrically, a clear understanding of narrative theories is important to the algorithmic designer. However, that understanding is not sufficient in order to build appropriate tools for Interactive Drama.

An Interactive Drama is composed of three components:

- the tool, which can be used by an artist to create an Interactive Drama (analogous to the camera in film making)
- the content, made of algorithms and data, which are entered into the tool by the artist through authoring interfaces (analogous to the film roll)

As depicted in Fig. 5, these three components are accessed by three different types of actors: the algorithmic designer, the artistic designer and the player. In this diagram is omitted the two practical skills: programming and artistic making, assuming they are possessed by the algorithmic designer and the artist). While only the player has access to the interactive experience, the other two namely the algorithmic designer and the artistic designer can interact with all the components enabling them to modify the tool and the content.

One of the main issues in the design process is the frontier between the tool and the content. The more complex and extensive the tool, the easier and the less creative the artist's work. For example, in the IDtension system, the tool does not come with a set of predefined features for characters. They must be defined by the author. This design choice was made because a feature system should not be imposed to the artist. Other systems for Interactive Drama, like Erasmatron [8] are different with this respect because they propose a predefined set of character's features. However, in the IDtension system a hard coded, non authored Narrative Logic is used [24], while it could have been decided to make it authorable. In this example, the choice between what should be part of the tool and what should be part of the content is hard to make. Defining the proper frontier would be a key outcome of the use of a co-design-based methodology for Interactive Drama.

Back to Fig. 5, the three types of persons involved in Interactive Drama (the algorithmic designer, the artistic designer, and the player) are traditionally dedicated to separate activities: the algorithmic designer makes a program, the artistic designer creates the work and the player plays with the work. The Interactive Drama system envisioned here would allow any designer to switch from one activity to the other. More precisely, it is the sequence of such activities which will constitute the basic design patterns of a methodology for Interactive Drama. Let us sketch some of these sequences:

---

Figure 4. The skills to build an Interactive Drama

Figure 5. Co-design of the tool and the content
• the test-based authoring sequence: The user has the possibility to write some content, then immediately play with it and tune his/her content accordingly. This is typically what is allowed by a good authoring tool: The author can write and test quickly, to improve the final result. For example, the project developed at the ZGDV in Digital Storytelling [21] has been designed to allow the user to quickly test the content, even if it is only partially developed. Another interesting approach where the user can tune a story in Virtual Reality while being immersed in the virtual environment has been explored in [26].

• the empirical modification of the tool (or tool): In that case, the user tests the story as a player and realizes the insufficiency of the models. S/he then modifies the tool itself to satisfy this new need.

• the "authorabilization" of the tool: after several tool modifications to accommodate new needs, the user decides that the corresponding parts of the tool could be made tunable by the artistic designer. The task of modifying the tool is replaced by a task of adding a specific part of content (in terms of rules or rules' parametrization).

• the tool-driven artistic creation: a specific characteristic of the tool (either a real feature or a limitation) generates new ideas for the content. This sequence from the tool to the content is fundamental, because it tends to produce stories which are specific to the medium, rather than mere adaptation from previous media.

These various sequences can of course be combined into more elaborate sequences. Furthermore, these sequences only represent part of the design activity: one should not ignore the fact that part of the design occurs outside the system itself, with pen and paper (and possibly with a typewriter or a word processor). The back and forth between the computer system and the pen and paper design is fundamental to Interactive Drama because of the abstract nature of the content: Even with the best authoring tool, the designer will always need to write fragments of stories without the constraints of the tool, because story creation naturally occurs at a concrete level [25]. Thus, a full methodology, which is only sketched here, must also include other sequences with more traditional writing tools.

The various design sequences described above break with the hierarchical and unidirectional relations between art and technology. Through a set of several interwoven design sequences, the work and the tool are created at the same time, through the activity of the designers. The tool and the content dynamically establish their functions and their frontier.

These dynamics suggest an analogy with a biological system which evolves and adapts to its environment. The system composed of the tool and the work is compared to an adaptive living organism. In Piagetian terms, the process of content modification is similar to the notion of assimilation (integration of a new skill/knowledge within an old schema), while the process of tool modification is similar to the process of accommodation (modification of an old schema to integrate new skill/knowledge). Applying this theory to the design of Interactive Drama is interesting because it acknowledges the status of the non-finished experimental work. Departing from the classical view of artistic creation, where the piece of art is the final result (aimed to be exhibited in a museum for example), the dynamic methodology is continuously producing "works in progress", which correspond to stages in the Piagetian theory of cognitive development. It is then up to the artist to decide which works in progress to exhibit. A possible end of the process occurs when the tool becomes stable, that is when no more accommodation seems to be needed in the short term.

It is also worth mentioning that the humans involved in this dynamic creation evolve and learn at different levels how to design a future Interactive Drama. This is a necessary condition, since creating an Interactive Drama requires a lot of progress from current artistic and technical practice. However, it could be detrimental in a certain way. If the first designers of the system become through learning and adaptation part of the system, then it is likely that the tool will accommodate (in part) only one specific artist. A way to overcome this limitation is to involve several artists in the loop, not necessarily during the complete production process, but at least at some periods in order to make sure that the architectural choices are meaningful for several artists.

Such a methodology is however difficult to implement in practice because of institutional rigidity. First, Art and Science are two different domains and two completely separate branches of the academic hierarchy of disciplines. Simply having an artistic designer and an algorithmic designer working together in the long term is difficult. Second, in the innovative labs which dare mixing skills, it is often the case that the collaboration is based on the idea of using technology/science to create art. But what is needed in Interactive Drama is the making of the technology/science and the art simultaneously, which is quite a different approach. Third, funding is made more difficult by the fact that creating an artistic work is not appealing to a Science lab while creating a technology is no more appealing to an Art Department or Institute. But there is a field where these limits have been overcome: Computer Music. We hope that in the future, Interactive Drama will be able to follow the path of Computer Music, by providing the ground for an effective integration of skills needed to create the future of Interactive Drama.

5. Conclusion
Interactive Drama has gained growing attention over the last years. It is now recognized as a difficult but not impossible challenge for Artificial Intelligence. In this paper, with our practical and theoretical experience in the field, we have tried to interpret the everyday difficulties encountered during our research in broader terms, by "looking at the future". Given the significant research done in Interactive Drama and connected fields in the last ten years, we do not claim that the main problem of the field is the lack of powerful algorithms to manage characters and plots. Rather, the future lies into a better coordination of the current material. First, powerful drama managers and believable agents have been developed so far, but their integration is primitive. We have proposed new avenues for this integration, which involve a more structured dialog between the two entities. Second, we have identified the need for new methodologies to provide both art pieces and tools for Interactive Drama. Such a methodology has been sketched and discussed.

This paper illustrates how long difficult and challenging the path leading to Interactive Drama remains. Presently, this path is not followed by the industry because it is too long and unpredictable. This is regrettable because the entertainment industry might be the best suited environment not only because of its available funding
but also by its skilled ability to manage projects with heterogeneous skills.

6. ACKNOWLEDGMENTS
This research is supported by an Australian Research Council Linkage International Fellowship Grant titled "An Interactive Drama Engine in Virtual Reality" (#LX560117).

7. REFERENCES