

Stepping into the Interactive Drama

Nicolas Szilas

LINC – University of Paris VIII
IUT de Montreuil
140, rue de la Nouvelle France
93100 Montreuil, France
n.szilas@iut.univ-paris8.fr

Abstract. Achieving a successful Interactive Drama where the user can act as a character in the story requires not only finding an algorithmic solution for combining interactivity and narrativity, but also interfacing those algorithms with the user. This paper focuses on the way in which the user can choose the actions of the character. Three specific issues are discussed: the variety of choices proposed to the user, the need for the user to anticipate his/her future possibilities for actions and the time necessary to enter the action. This allows us to propose a taxonomy of different user interfaces and to evaluate the advantages and drawbacks of each category of interface. This should serve as a guideline for the design of user interfaces for Interactive Drama.

1 Introduction

This paper addresses the specific kind of interactive dramatic experience on the computer, where the user is a character in the drama. In short, we call this an Interactive Drama (ID), even if this term sometimes covers a larger set of experiences. Even if the idea of building an ID has arose for some years now [3,9,13] and despite some research effort on the subject [4,11,15,21,22,23], ID appears to be a quite difficult issue, which requires the collaboration of various disciplines.

Within the *IDtension* project [8], we study ID from three different but interrelated approaches [18]: the algorithmic approach [19,20,21], the author approach[18] and the user approach. This paper focuses on this last approach. In particular, it aims at providing some guidelines for the design of how the user enters his/her actions.

This topic has been rarely tackled even if several systems propose their own method for entering the action. This is certainly due to the fact that this issue is seen as a design issue, only involving an ergonomic design of the user interface. However, we believe that this issue is far more fundamental and that it is linked to the basic properties of ID.

This paper has two main goals:

- help the design of an ID;
- provide a set of conceptual tools to better analyze interfaces used in ID systems or related systems (like video games).

This paper is divided into three sections, each one focuses on a specific feature of ID. From the careful examination of each feature, we draw some conclusions regarding

the way to enter actions in an ID. We do not consider that we have exhausted the subject, other features of ID will be added later to this study.

2 The variety of choices

2.1 The choice problem

The most notable difference between ID and existing forms of interactive narrative (hypertext, Interactive Fiction, adventure video games, etc.) is the number of narrative actions that the user can undertake (the range of actions, to use Brenda Laurel's terminology [9 p. 20]). In an adventure video game for example, only a few actions really have a significant effect on the story (usually, only one action makes the story go forward, others are "fails"). This difference is particularly marked in the dialogs: being that a key feature of interactive drama is the dialogs, ID allows a user to have a wide range of dialog choices during the interaction [5,17,19]. For example, if one considers only one type of dialog act such as "ask for assistance", suppose that the user's character has to perform 5 tasks and that there are 5 other characters, it makes 25 possible "ask for assistance" acts. In *IDtension*, there are many types of acts. We could observe experimentally the fast growing number of choices given to the user [19].

Choosing among a large number of actions is a problem. Using a choice list of actions is obviously unsatisfactory if the number of actions exceeds 10. We call this problem the "choice problem".

In the example above, it is assumed that the algorithms for ID have the possibility to interpret all the actions and produce the expected effect. As a consequence, by giving all these meaningful choices, such an ID provides agency, as defined by J. Murray: "the satisfying power to take meaningful action and see the results of our decisions and choices" [13, p. 126]. The "choice problem" can be seen as the other side of the coin of agency in ID.

2.2 The interface mapping function

In order to properly classify the various types of interfaces that can be proposed to cope with this problem, we introduce the following notions:

- when the user is to choose an action, there is a set of actions that the system can logically process as a significant input to its narrative engine. Let us call this set *L*.
- at the same time, there is a set of actions that the user can physically perform and which she/he can interpret as significant from a narrative point of view. Let us call this set *P*. For example, a simple move of the character is not part of this set, while revealing an intention to another character would be part of this set.

- Let us define the ideal interface mapping function f the relation from the set P to the set L , which associate whenever possible the physical action to the proper logical action (see Fig. 1). By “proper” we mean that the interpretation by the machine of $f(x)$ is the same as the interpretation of x by the user, x being an element of P .
- Let us define the actual mapping function g the relation from the set P to the set L , which is implemented by the system and which associate, whenever possible, a physical action to a logical action.

The difference between f and g lies in the fact that in some cases, the theoretical function f is not easily programmed on a computer. Initially, we will consider that f and g are similar and we will only reason on f .

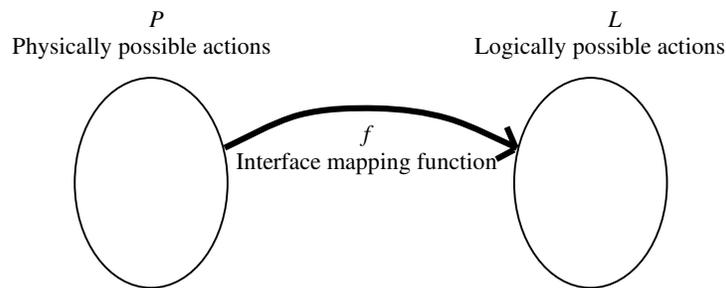


Fig. 1. The mapping function

The distinction between P and L looks similar to the distinction between perceived affordances and real affordances in ecological psychology and Human-Computer Interaction. However, we are not concerned with the immediate possibility of action like moving within the virtual world but with the higher level narrative actions, and the corresponding affordances, usually involving language. Higher level affordances have been discussed in the context of ID in terms of material and formal affordances [10]. While material affordances correspond to the set P, formal affordances, as defined by Mateas, are different from the set L, because they are linked to the *motivation* to act (“why [players] should take action within the story world *at all*”).

Let us specify that f and g are effectively functions (an element cannot be in relation with two elements) because a physical action is always interpreted univocally by the computer.

Depending on the simple mathematical properties of f , one obtains various types of interfaces for an ID, as illustrated in Fig. 2.

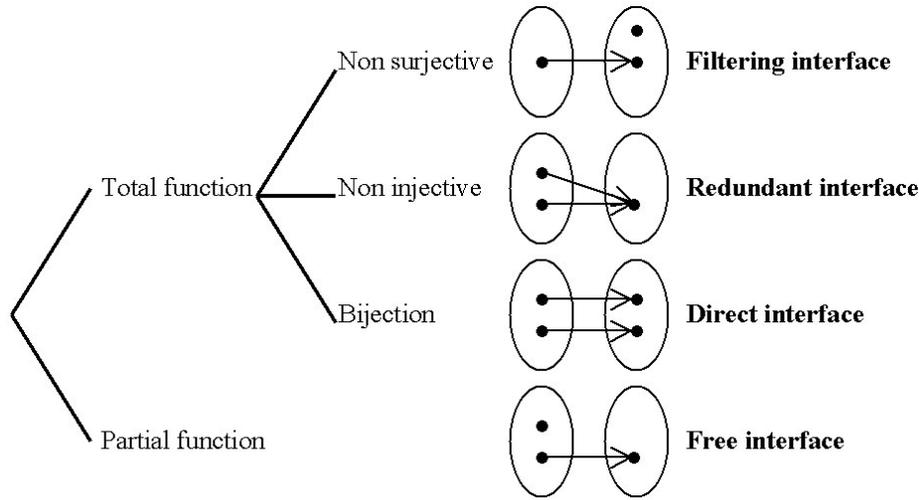


Fig. 2. Taxonomy of various interfaces for ID depending of the properties of the interface mapping function.

If f is not total, that is if some physical actions cannot be associated to logical actions, then we have a *free interface*. Indeed, the user is free to perform some actions, which are not interpreted by the computer.

If f is not surjective, that is if some logical actions cannot be reached by any physical action, then we have a *filtering interface*. Indeed, the interface acts as a filter, forbidding some logical actions to be chosen by the user.

If f is not injective, that is if two different physical actions are associated to the same logical action, then we have a *redundant interface*.

Finally, if f is bijective (injective and surjective), that is if there is a one-to-one mapping between the physical actions and the logical actions, then we have a *direct interface*. Indeed, the proposed actions are exactly the logical actions.

In the following, we discuss in detail each of these categories of interfaces.

2.3 f is a total function

If f is total, three cases are worth studying: the non surjective case, the non injective case and the bijective case (see above).

The filtering interface hides the complexity of the set L by only providing a limited set of options. Typically, it corresponds to the use of interface agents such as “wizards” or anthropomorphic agents. For example, in an ID the filtering interface consists in proposing to the user the choice of some of the most relevant actions even if the system could interpret more actions. The filtering interface solves the choice problem, because the size of L is hidden. This kind of interface, however, is problematic when the user wants to perform an action, which is not proposed. In that case, frustration takes place. We will discuss again this kind of interfaces in Section 3.

The redundant interface is classical in the design of user interfaces. However, in the case of ID, it does not solve our choice problem, because the redundancy increases the number of choices rather than decreasing it. Note that some free interfaces can have some redundancy, but it is not labeled a redundant interface in our taxonomy.

The direct interface consists in letting the user choose among a large selection of actions, through a well designed interface. Typically, real-time strategy games like *Warcraft*TM or simulation games like *The Sims*TM use that kind of interfaces, even if these games are not ID. The advantage of the direct interface over the free interface is that it solves the frustration problem of the free interface (see below). However, it does not solve the choice problem a priori, because if L is large, P is also large. Thus, the direct interface relies on the feasibility of an easy-to-use interface, which would allow the user to ergonomically select an action among tens of possible actions. Furthermore, this kind of interface, which is based on a complex combination of classical user interface components (lists, buttons, etc.) tends to disrupt the feeling of immersion.

2.4 f is a partial function

The free interface consists in giving to the user more expressive power than what the machine can really interpret. Typically, this interface uses a free text or free speech interface for entering the action. This type of interface has been chosen by several research projects on ID [5,11,16].

Among free interfaces are *natural interfaces*, which use free speech, free text, body gesture, etc. and *verb-based interfaces* (there might exist other types of free interfaces that we omit or forget to mention). Verb-based interfaces are used in some adventure video games (like the *Indiana Jones* series from *LucasArt*): the interface contains a panel with several clickable verbs: to choose the action, the user clicks on a verb, then on a point on the screen. For example, the user would click on “Push” and click on an object, to move it. These verb-based interfaces are free interfaces because the user could produce many more actions that the game could interpret.

The main advantage of the free interface is that it allows the user to easily choose among a large set of actions since the interface is closer to the human means of expression (especially for natural interfaces).

The problem of the free interface is that the user can choose an action that the machine cannot take into account. One can only hope that in the future with the improvement of narrative engines the set L will become larger, however, it will never cover the range of a free text or a free speech interface. Note that this problem is independent of the limitation of the language technology used to decode the user input (it remains even if f and g are identical).

With the free interface, the user has to control and limit his/her language to what the machine understands. The question is whether the user will manage to do this naturally or if this will provoke a feeling of frustration. This is related to the design of the free interface, whether it affords actions in P that are in L or not. By proposing two explicitly choices which are not in L , the verb-based interface typically generates bad perceived affordances.

Note that the classical turnaround of this problem is to add a special element in L which means “I don’t understand”, and which encourages the user to change or rephrase his/her action. While video games have used this technique in a basic way, there exists a more subtle way to do it which consists in having numerous general reactions to a non interpreted event [17].

It is formally possible for free interfaces also to distinguish the non-injective and non-surjective cases. Practically, the free interfaces aim to be surjective (no filtering): it is their *raison d’être* (even if g could be non-surjective, for technical problems or limitations). The natural interfaces are redundant (f is not injective) while the verb-based interfaces are not (f is injective).

2.5 Conclusion on the taxonomy

Until now, we have only considered a single mapping function f , hence a single type of interface. However, it is conceivable that a more complex interface could combine several modes of interaction. To each mode is associated a set P and a mapping function f . For example, the user is proposed a small set of actions, (filtering interface) if no action is satisfactory, he/she will switch to a text field to type another action (natural free interface). Such “hybrid” interfaces might be interesting to explore in order to combine some advantages of the various types of interfaces mentioned above.

The rigorous study of the mapping function has allowed us to classify the interfaces regarding how they tackle the choice problem. However, we need a deeper analysis to better understand the differences and limitations of those various types of interfaces. Some elements of such an analysis are proposed in the next section.

3 The anticipation of action

3.1 Anticipation in classical and interactive narratives

The force of any narrative lies in the way the audience’s expectations are handled. Umberto Eco’s “inferential walks” refer to the various paths the reader imagines at a given point of a narrative [6]. At each path is associated a subjective judgment of possibility. The author activity consists in planning the reader’s inferences in order to maintain his/her interest. The reader tries to anticipate the evolution of the story, making good and bad guesses, which provides for an enjoyable and playful experience.

Narrative effects like suspense or surprise are directly related to this play. These inferences are performed according to some sort of rules, most of which do not show in the narrative itself: the reader uses his/her own encyclopedia, which contains all his/her knowledge of the real world, on the genre and on other narratives [6]. Thus, dur-

ing the narrative, the audience is constantly making anticipations on what could happen, both immediately and on a longer period of time.

In ID, where the narrative becomes interactive, we estimate that this essential feature of narrative must be preserved. This means that the user must be able to anticipate events and actions including his/her own actions because the user is a character. Because the user's actions are mediated through the interface, this necessity of anticipation has consequences on the design and understanding of the interface.

The rules used by the user to anticipate what he/she will be able to do later are contained not only in the "encyclopedia" of the user but also in the "rules of the game", which are set before or during the interactive drama. For example, in *GTA III*, the user can anticipate that his/her character will steal a car later in the game. However, he/she cannot anticipate that the character will invite passengers in the car because she has internalized the possibilities and impossibilities in the game at a given moment in the play.

3.2 Consequence on the filtering interfaces

It is necessary that the user anticipate which possibilities he/she will be given later in the game. Given this principle, some interfaces mentioned above appear not to be suitable to ID. Indeed, filtering interfaces where only a subset of logically possible actions are presented to the user make the anticipation of action impossible. The system chooses a limited set of possible actions and the user cannot anticipate which actions will be possible, and which ones will not be (the user cannot anticipate the content of the set P , even if L is predictable). Typically, in many adventure video games, the user has to choose among a small set of possibilities. In this situation, he/she discovers these possibilities as they appear on the screen: he/she could have anticipated these possibilities or not. More importantly, the user could have planned to perform an action but this particular action would have not been proposed as a choice. This makes the adventure game genre a narrative genre where the user is still in the position of a person to whom the story is being told. He/she is being given some successive choices in the narrative, which are sequentially defined by the author. Being unable to anticipate the choices, the user is not capable of building a kind of strategy: he/she is no longer a character. Note that this critic on adventure video games is independent of the limited algorithmic possibilities of these games, in terms of management of user choices: this is an interface problem that would occur in an AI-based ID that would make use of a filtering interface.

3.3 About stability...

In this discussion, we have argued that the user should be able to anticipate the set P of physically possible actions. This is also true for the set L . This means that the set of possible actions must be relatively stable during the narrative.

It also means that the user must perceive this stability. For the direct interface (see previous section), this stability is perceived through the stability of the user interface itself. This task is more difficult when it comes to the free interface: it is central for

this kind of interfaces that the user expresses him/herself within the proper range, so that his/her actions are understood. After an initial period of trial and errors combined with a playful behavior of testing the limits of the systems, the user has to implicitly understand the size of the set L . From an authoring point of view, it is necessary for the author to accurately guess the possible behaviors of the user in order to design the set L . We find here a process similar to the classical authoring of text as described in [6]. Instead of predicting internal inferential paths (classical story) the author predicts an actual action of the user. This is more critical because a bad prediction can lead to a failed action and to corresponding frustration.

The anticipation of action also has an impact on short-term interaction. When the user has to act, if he/she already knows (has anticipated) the possible actions then the interaction becomes user friendly. For example, in a filtering interface again, the user has to read the proposed choices, whereas in other interfaces because of their stability the user already has his/her internal representations of the set of possible actions.

3.4 And surprise!

The argument of stability is in contradiction with the idea of surprise: if the user can expect all possibilities of action, he/she cannot be surprised by a new way of acting (he/she can still be surprised by the outcomes of the actions, actions of others characters, and by events). That is why we propose to weaken the stability constraint by allowing the inclusion of new possibilities of actions, at any given moment in the narrative. This is typical in strategic and adventure games: new "powers", new tools, new buildings are given to the user to act, and sometimes the user does not expect these new possibilities. Two important constraints must, however, be followed in order to maintain the anticipation of action:

- once the new possibility to act is added, it should remain (which is the case in the video games);
- the pace of the adding of new and unexpected possibilities of action must be slow.

This second condition is necessary so that these new actions are not taken into account in the "inferential walks" of the users. If new possibilities of action appeared often, then the user would implicitly tell him/herself "If I do this, may be I will be given new possibilities, maybe not"... which is disturbing.

*

To sum up, the extension of the fundamental narrative notion of anticipation to the field of interactive narrative has allowed us to better analyze and understand various user interfaces for ID. Interestingly, this study also shed a new light on some existing game interfaces.

4 The duration of interaction

4.1 The problem

Another fundamental difference between many current real-time 3D video games and ID is the amount of time given to the user to act. While in real-time 3D video games the fundamental action is to move the character through the virtual space, in ID, the idea is to *think* for the character. The time scale is different between the two experiences: in action games, the user reacts in a tenth of a second, while in ID, several seconds are available to act.

This amount of time is useful for two reasons:

- some dramatic situations are complex and ask for a certain reasoning, especially for conflicting situations;
- the user interface can be complex, especially if a direct interface is chosen.

In ID, the time scale to act is thus longer than the time scale of the virtual world. Emblematic to this problem is the so called *waiting animation* in video games, where the character is having some various futile gestures either realistic (looking around), or humorous (playing basketball with its body, *Rayman*TM).

It could be argued that this is a transitory problem due to the limitations of the current interfacing technologies. With a perfect speech and body gesture recognition system, the user would be able to interact and think continuously as he/she does in real life. However, we still consider that the problem of the duration of interaction must be seriously considered, for the following reasons:

- currently, we do not have such "perfect technology", yet we wish to produce some good examples of ID;
- discarding the problem returns to discard direct interfaces, which have some advantages to be explored (see Section 2);
- it might be stressful to ask the user to make all the choices in real-time and we want to explore the case where the user does have time to think of his/her actions.

4.2 Classical solutions

In order to compensate this difference of scale, there are two classical solutions:

- freeze the fictional time during the interaction.
- Fill in the time to act by a neutral activity, that is an activity that has no significant influence on the narrative.

The first solution has the drawback to disrupt the immersion.

The second solution consists in having some waiting animations as mentioned above. It works up to a certain point but it is obviously unnatural to have your character looking around for 5 or 10 seconds...

4.3 The semi-autonomy

Another solution consists in letting the user's character take actions. The character is then semi-autonomous [14]. This solution, which amends the very definition of ID presents several difficulties because the user is not fully responsible for his/her character:

- where to draw the limit between autonomous and controlled behaviors?
- would the user appreciate the lack of control over his/her character's behavior?
- how to design the communication between the sometimes autonomous character and the user?

Thus, the solution of semi-autonomous is interesting to "fill in" the duration of interaction but it is certainly quite difficult to design.

4.4 From immersive to elliptic ID

There is a last solution that we would like to consider here. It consists in freezing the time during interaction as mentioned above but not necessarily restarting the fictional time exactly where it stopped. For example, the user would choose to "leave the house", and after entering his/her action, the character would be in the garden. This mechanism is well known in various narrative forms and is called an *ellipsis*. The idea here is not to undergo the interruption caused by the duration of interaction but conversely to take advantage of it by transforming it into an ellipsis.

Such use of an ellipsis in ID is certainly a breach of the classical vision of ID as it has been conveyed by J. Murray in [13]. Indeed, this vision considers Virtual Reality as the ideal medium for ID, and Virtual Reality is fundamentally real-time, a characteristic that must be taken into account in order to design an ID for this media [2]. In a more general approach to ID, however, we must consider the computer as a system for Virtual Reality applications and other applications. In that context the use of the ellipsis in ID should be taken into consideration.

The classical theories of narrative consider two times: the narrative or discursive time, which corresponds to the time of the narration process, and the story time, which correspond to the time of the world of the story (diegesis) [7]. An ellipsis is a typical relation of duration between those two times (other relations being a descriptive pause, a summary, etc.). The ID introduces very naturally a third time, the time of (inter) action, which is a sub-category of the narrative time. In other forms of interactive narratives, like hypertext, this third time has less importance because it does not conflict with other times. But in ID, because of the very nature of drama which consists in showing characters acting (the mimesis [1]), the duration of interaction is to be taken into account. In drama, within a scene, the narrative time and the story time run in parallel. In ID, the introduction of the interaction time disrupts this parallel evolution and possibly injects ellipses at the heart of drama.

There exists a well-known form of drama (drama being defined as a mimetic narrative form) which uses ellipses intensively: comics. In comics, the interstice between two boxes usually represents a time ellipsis between the events described in the boxes. This suggests that while cinema is the most common media of reference in ID

[2], some interesting forms of ID could also derive from comics. It does not mean that it would be made of fixed images (although this option is to be considered) but that some forms of ID might also be some “sequential art”, as defined in [12].

5 Conclusion

Starting from features specific to ID we have proposed a taxonomy of input interfaces for ID and raised a certain number of issues regarding the design of the interaction.

We have tried to be objective and not to bias our study towards a certain type of interface. Practically, we conclude that the viable solutions are the following:

- the free interface (if the technical issue and the user frustration issue are reasonably solved);
- the direct interface (if a user-friendly interface is found, and the disruption of immersion made acceptable);
- the filtering interface, only in combination with one of the above interfaces.

For the *IDtension* engine [8], we have chosen a direct interface and we are currently designing the interface.

Interestingly, the specific issue of the duration of interaction in the virtual world has led us to propose alternative forms of ID, which would be inspired more by comics than movies or theatre. By their elliptic nature these forms challenge in someway the widespread vision of ID, the “Holodeck vision” [13].

References

1. Aristotle, 330 BC. *La Poétique (The Poetics)*
2. Aylett, R., Louchart, S. Towards a narrative theory of Virtual Reality. *Virtual Reality*, 7(1), 2-9, Dec. 2003
3. Bates, J. Virtual Reality, Art, and Entertainment. In *Presence: The Journal of Teleoperators and Virtual Environments*. Vol. 1. No. 1. MIT Press. (Winter 1992)
4. Cavazza, M., Charles, F., Mead, S. J.: Characters in Search of an author: AI-based Virtual Storytelling. In *Proceedings of the First International Conference on Virtual Storytelling (ICVS 2001)*. Lecture Notes in Computer Science 2197, Springer Verlag (2001) 145-154
5. Cavazza, M., Martin, O., Charles, F., Mead, S. J. Marichal, X. User Acting in Mixed Reality Interactive Storytelling. In *Proceedings of the Second International Conference on Virtual Storytelling (ICVS 2003)*, Toulouse, France, Lecture Notes in Computer Science, n. 2897, Springer Verlag, 189-197 (2003)
6. Eco, U.: *Lector in Fabula.* Bompiani, Milano (1979)
7. Genette, G. *Figure III*. Paris, Le Seuil (1972)
8. IDtension. <http://www.idtension.com>
9. Laurel, B.: *Computers as Theatre*. Addison-Wesley, Reading Harlow Menlo Park Berkeley Don Mills Sydney Amsterdam Tokyo Mexico (1993)
10. Mateas, M. A preliminary Poetics for Interactive Drama and Games. *Digital Creativity*, 12 (3), 2001, 140-152

11. Mateas, M., Stern, A.: Towards Integrating Plots and Characters for Interactive Drama. in Proc. AAAI Fall Symposium on Socially Intelligent Agents: The Human in the Loop (North Falmouth MA, November 2000), AAAI Press (2000)
12. McCloud S. Understanding Comics: The invisible Art. HarperCollins, Publishers, Inc., New York (1993)
13. Murray J. Hamlet on the Holodeck. The future of narrative in the cyberspace. Free Press, New York (1997)
14. Portugal, J.-N. Environnement narratif: une approche pour la fiction interactive, appliquée au jeu The Insider. Imagina'99 (1999)
15. Sgouros, N. M. Dynamic, User-Centered Resolution in Interactive Stories. In *Proc. IJCAI'97* (1997)
16. Spierling, U., I. Iurgel. "Just Talking about Art" – Creating Virtual Storytelling Experiences in Mixed Reality. In Proceedings of the Second International Conference on Virtual Storytelling (ICVS 2003), Toulouse, France, Lecture Notes in Computer Science, n. 2897, Springer Verlag, 179-188 (2003).
17. Stern A., Mateas M.: Integrating Plot, Character and Natural Language Processing in the Interactive Drama Façade. In Göbel et al. (eds) Proc. TIDSE'03. Fraunhofer IRB Verlag, (2003)
18. Szilas, N. , Marty, O., Rety, J.-H... Authoring Highly Generative Interactive Drama. In Proceedings of the Second International Conference on Virtual Storytelling (ICVS 2003), Toulouse, France, Lecture Notes in Computer Science, n. 2897, Springer Verlag, 37-46 (2003)
19. Szilas, N.: IDtension: a narrative engine for Interactive Drama. In Göbel et al. (eds) Proc. TIDSE'03. Fraunhofer IRB Verlag, (2003)
20. Szilas, N.: A New Approach to Interactive Drama: From Intelligent Characters to an Intelligent Virtual Narrator. In Proc. of the Spring Symposium on Artificial Intelligence and Interactive Entertainment (Stanford CA, March 2001), AAAI Press, 72-76 (2001)
21. Szilas, N.: Interactive Drama on Computer: Beyond Linear Narrative. In Papers from the AAAI Fall Symposium on Narrative Intelligence, Technical Report FS-99-01. AAAI, Press Menlo Park (1999) 150-156
22. Weyhrauch, P. Guiding Interactive Drama. Ph.D. Dissertation, Tech report CMUCS-97-109, Carnegie Mellon University (1997)
23. Young, R.M.: Notes on the Use of Plan Structure in the Creation of Interactive Plot. In Papers from the AAAI Fall Symposium on Narrative Intelligence, Technical Report FS-99-01. AAAI, Press Menlo Park (1999) 164-167