

Thinning the Fourth Wall with Intelligent Prompt

Rossana Damiano^{1,3}, Vincenzo Lombardo^{1,3}, and Antonio Pizzo^{2,3}

¹ Dipartimento di Informatica, Università di Torino

² Dipartimento di Studi Umanistici, Università di Torino

³ CIRMA, Università di Torino

Abstract. This paper presents a digitally enhanced model of performer–audience communion in an interactive storytelling setting, based on an intelligent prompt system. The audience response is taken into account through emotion detection; the performer decides about her/his attitude towards the audience. The intelligent prompt advises the performer about how to continue the story. The model, named DoPPioGioco (“DoublePlay”), has been implemented as a prototype system, a virtual environment that realizes a feedback loop between the performer and audience.

Keywords: Intelligent Prompt, Story Editing, Emotion Models

1 Introduction

This paper describes a model, called DoPPioGioco (“DoublePlay”), for permeating the fourth wall, “the age-old dream” of interactive storytelling [8, p. 1]. The model assumes a performer–audience relationship mediated by a system, the “intelligent prompt”, which automatically prompts the next story chunk to be acted out by the performer in order to properly reacts to the audiences emotional engagement. DoppioGioco assumes a core emotional model that is employed for tagging the story chunks and classifying the audience reaction, and a real time engine that prompts the next chunk to be delivered as a consequence of the audience’s response and the performer’s current attitude. In this paper, we describe the model and a prototype system built for teaching interactive storytelling to a class of media and drama, eventually reflecting on the results of the workshop.

The inspiration for DoPPioGioco draws from two main sources. On the one hand, the line of research in studying and designing the experience of the audience, situated at the junction of HCI and media studies, surveyed by Brooker [6]. In particular, Tanenbaum reconsiders the notion of “subversive player” [19], developed in game studies to describe the user’s attempts at breaking the boundaries of the interactivity allowed for by games, with the goal of making this behavior an intrinsic propellant of the game. On the other hand, the paradigm of improvisational theatre, which has opened the way to a computational approach for dominating the complexity of emergent storytelling [18, 2]. The dynamics of improvisational theater have been described by Baumer and Magerko [5] in the

perspective of interactive storytelling, using the Decision Cycle from Newell’s Unified Theory of Cognition (*receive* new inputs, *elaborate* new knowledge, *propose* actions to take, *select* one of those actions, *execute* the action) as a conceptual framework for analyzing the way each performer takes advantage of the scene advancing moves of the others.

Drawing inspiration from the performers’ ability to interpret the script in response to the audience reaction [1], DoPPioGioco acknowledges the traditional distinction between the roles of the author and the performer in story design and delivery. The narrative component is mostly handled offline, at the story editing time, as a combinatorial explosion of the possibilities arising from the storyline and emotional tagging system. DoPPioGioco is implemented as an interactive storytelling platform that offers the author an environment for crafting the story as a graph of story chunks and a function that prompts the performer by interpreting her/his task with an emotional rather than narrative matrix. The core of the intelligent prompt is the emotional model: emotions tag the story chunks, emotions are detected from the audience reactions, emotions are employed to prompt the next chunk.

2 Emotional Systems in Interactive Storytelling

Emotions have received much attention in interactive storytelling, both as a component of artificial character models, and as a feature of the story plot. While in synthetic characters emotions have become part of the processes that generate the character’s behavior [7], thanks to the adoption of specialized models of emotional appraisal and coping issued by psychology [3], the role of emotions in the interactive generation of the plot has focussed on the engagement curves of Façade [11] and the emotional trajectories of emergent narrative (see the historical perspective in [10]). In the interactive drama Façade, the generation of the story was driven by a function that kept the emotional engagement of the user close to a target curve; in the Distributed Dramatic Management of the Emergent Narrative [20], predetermined emotional trajectories were employed as a metric for dramatic impact in the selection of characters’ actions. With the exception of the systems mentioned above (see also [21]), most interactive storytelling systems do not directly address the emotional engagement of the audience. Rather, the research in interactive story generation tends to focus on the consistency of the plot in terms of characters’ intentions and actions (see for example [13]). A notable exception is given by the use of player modeling to affect the emotional response of the users in [14].

Given the requirement of appraisal with respect to story chunks delivered through some audiovisual medium by the intelligent prompt, DoPPioGioco relies on the GEMEP model [4]: originally designed to support the creation of a corpus of clips displaying the performance of emotions by human actors, the GEMEP model is based on an extensive survey of the existing theories and models of emotions, including cognitive and dimensional models. GEMEP, GENEVA Multimodal Emotion Portrayals, is a collection of audio and video recordings

featuring actors portraying affective states, with different verbal contents and different modes of expression. Thanks to its syncretic and methodologically robust design, geared on performance, this model is especially suitable to annotate the affective content of media. In GEMEP, emotions are grouped along two axes: the *polarity* (positive/negative) and the *intensity* (high/low). The combination of these two axes provides four emotion families, each including three emotion types; within each family, the emotion categories are mainly characterized by different arousal levels:

- **Positive, high intensity:** amusement, pride, joy;
- **Positive, low intensity:** relief, interest, pleasure;
- **Negative, high intensity:** hot anger, panic fear, despair;
- **Negative, low intensity:** irritation, anxiety, sadness.

In the following, we introduce the DoPPioGioco model and the implemented system. Then, we describe the case study and the lesson learned from the experimental usage of the prototype as part of the teaching methodology.

3 Model DoPPioGioco

Basically, the DoPPioGioco system consists of an intelligent prompt, which navigates a story chunk graph annotated with emotions, and handles the performer’s attitude towards the detected emotional response of the audience. The design of the graph of pre-authored story chunks follows a model of exhaustive combinatorics with respect to the four emotion families, with each story chunk annotated with a set of emotions. The performer can choose her/his attitude over two behavioral axes, the pleasant/opponent axis, and the low/high arousal axis, respectively (notice that both axes address the polarities of the GEMEP model). At each step, the audience responds to the last played story chunk; then, the initiative goes to the performer: she/he decides whether to empathize with the audience (*pleasant*) or to antagonize it (*opponent*), and whether to respond with a *low* or *high* arousal. Then, the intelligent prompt selects the next chunk among the suitable ones, depending on the interplay of the audience’s emotional reaction and the selected attitude (one out of four combinations: *pleasant/low*, *pleasant/high*, *opponent/low*, *opponent/high*).

DoPPioGioco can be described through its two main functions, exemplified by the two system interfaces: the offline Story Manager editor, which organizes the story chunks in a story graph and supports the tagging of the story chunks (*units* in DoppioGioco terms); the online Stage Manager, which implements the intelligent prompt function by managing the story advancement in response to the audience’s reactions.

The Story Manager (Figure 1) allows the author to create and organize the units offline. Each unit consists of an audiovisual clip and a set of metadata elements describing it, such as title and textual. For each unit, the author has to provide the information needed to the story engine to create a consistent story at runtime: the precedence relations with the other units, needed to generate



Fig. 1. DoPPioGioco. The interface of the Story Manager.

a causally motivated story, and the emotions attached to the unit, needed to account for the response of the audience. Each unit must be tagged with a set of emotions to ensure that the story engine can drive the story advancement consistently based on the combination of the emotional response of the audience and the attitude chosen by the user, thus enabling the intelligent prompt to break, metaphorically, the “fourth wall” (technology enhances the connection between the performer and the audience). Figure 1 shows the interface for the editing of the units. The interface is divided into four areas: the top left area contains the unit metadata; the top right area allows uploading and viewing the media asset that realizes the unit (typically, an audiovisual clip); the bottom left area contains the emotion tags: for example, the unit displayed in the figure is tagged with the emotion “sadness”; finally, the bottom right area allows specifying the units that precede and follow the unit. The story chart editing tool (see Figure 3 in the next section) allows the author to manipulate the position of the units in the story chart.

When the system is run in the online mode (the intelligent prompt “on stage”), the Stage Manager enters into play (Figure 2). After each unit, the audience responds by displaying a set of emotions: the individuals react in different ways, so the system computes the most frequent emotion, possibly raising to the family level (in case of wide variety). At this point, the performer can decide to *please* or *oppose* the audience, and set the intensity to *high* or *low*. After setting the reaction type, the system picks up the next unit among the continuations allowed by the story graph. If several continuations are compatible with the current story unit and the response of the audience, the system selects one randomly. Figure 2 shows the interface for managing the stage. The interface is horizontally divided into two areas which represent, respectively, the performer’s actions (top area) and the audience’s response (bottom area). The

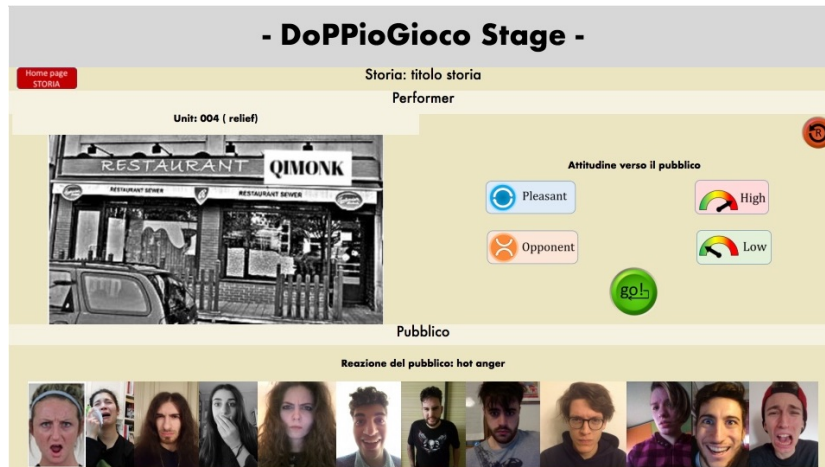


Fig. 2. DoPPioGioco. The interface of the Stage Manager.

top area contains, on the left, the last played unit; the right area is occupied by the performer’s console for managing her/his attitude in the reaction to the audience: the four buttons refer to the global parameters of the emotional control of the continuation, namely polarity (negative or positive, i.e. “pleasant” or “opponent”) and intensity (“low” or “high”) – remember from Section 4.2 that each combination corresponds to an emotion family. The lower part of the interface contains a button for getting the emotional response of the audience (“Get response”, not visible in the figure): when pressed, this button reveals the reaction of the audience and the button becomes invisible until the next unit is selected. For attaining a more realistic impression, the facial expressions of the audience members are displayed (in the figure, the by-a-majority emotion of “hot anger” has been computed by the system for the audience). The “go!” button posited below the console (in green, see Figure 2) triggers the selection of the next unit: the new unit is loaded, the console is reset and the faces of the audience are replaced by the button “Get response”.

4 The DoPPioGioco Implemented System

In this section, we illustrate the architecture of the system, describing the role played by the emotion model from the annotation of the units to the intelligent prompt.

4.1 Narrative-Emotional Model

Following an established practice in interactive storytelling, the story is formally structured as a directed graph (see an example from the case study in Figure 3). The interaction with the audience determines the transition to the next node:

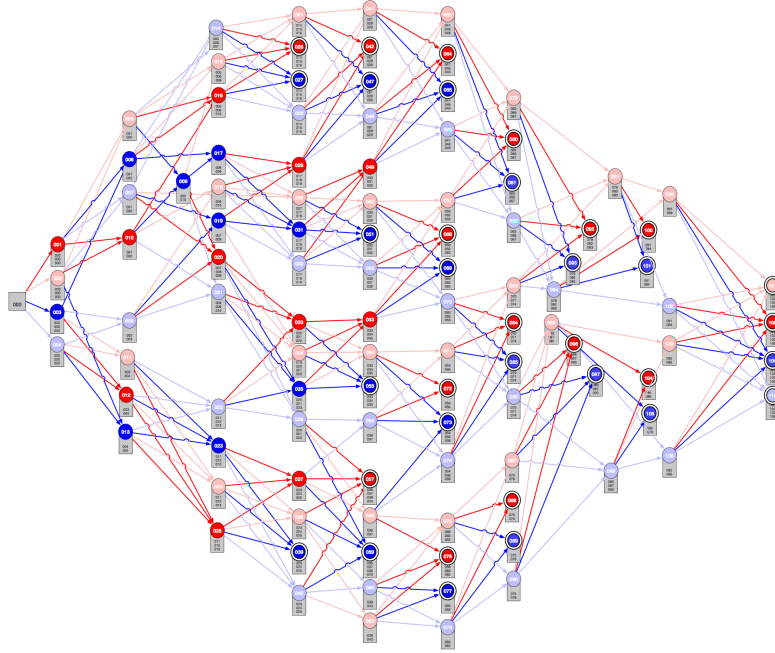


Fig. 3. The visualization of the story plot.

in DoPPioGioco, the transitions that are not compatible with the response of the audience are filtered out by the emotion model superimposed to the story graph, and the decision about the continuation is carried out by the intelligent prompt described above.

This approach requires that the story units are annotated offline, by encoding not only their position in the story graph, but also the affective information brought by each story unit according to the author. So, each unit is labeled with a set of emotion tags (see Figure 1), which represent the emotions that the author expects the audience to feel when the unit is delivered – based on its narrative content. The tags correspond to the emotion types encompassed in the GEMEP model (see Section 2). Finally, each unit must be associated with a video clip, that provides indications on what and how to perform ‘on stage’, when the unit is selected.

4.2 Annotating Emotions

In DoPPioGioco, the story units are annotated with the 12 emotion categories of the GEMEP model.⁴ Issued from a compromise between the dimensional [12,

⁴ The GEMEP model also encompasses 5 extra emotion types that don’t fit in the categorial classification (Admiration, Tenderness, Disgust, Contempt, Surprise); they have been omitted from the emotion annotation in DoPPioGioco.

15] and the cognitive models of emotions [17], the polarity-based account of emotions incorporated in GEMEP is suitable to deal with the polarity of the reaction to the audience’s response implied by DoPPioGioco: as anticipated in Section 4.2, the the decision to play against (or pro) the audience can be directly mapped onto the negative/positive dimension of emotions in GEMEP, with “against” corresponding to the “opposite polarity” and “pro” to “same polarity” (see the definition of the emotion families in the previous section). The following *Reaction rules*, applied in a cascading way, determine the continuation of the story after collecting the audience’s response and the performer’s subsequent reaction:

- **R1:** *If the selected choice is **pleasant**, then select the emotion families with the same polarity; else (the storyteller decides to be **opponent**), select the families with opposite polarity.*
- **R2:** *Given the polarity established by R1, tune the intensity level of the reaction to the selected intensity (low or high).*

In the implemented system, in order to emphasize the elements of arbitrariness that characterize a live, interactive performance, a random element was introduced in the selection of the emotions within the selected family: given the available units, the system randomly selects the next unit among the available ones, so that the user does not have complete control on the selection.

For example, consider the unit represented in Figure 1. The unit, 058, is tagged with “sadness” (Negative, Low intensity). Let’s assume that the audience responds with a negative emotion, as expected (for example, “despair”, which belongs to the Negative, High Intensity family). In this case, if the performer decides to play in favor of the audience and to keep the same intensity level, the system will select, among the possible continuations, the units tagged with the same emotion family as the emotion expressed by the audience; so, a unit tagged with “irritation”, “anxiety” or “sadness” may be chosen. On the contrary, if the performer chooses to play against, and with a high intensity, the rules would select only units belonging to the Positive, High Intensity family (“amusement”, “pride” and “joy”).

4.3 System Architecture

The architecture of the system encompasses four main modules (see Figure 4): the Story Manager, the Stage Manager, the Emotion Manager, and the Audience Manager. The knowledge about the story and the media assets, created and uploaded offline through the interface displayed in Figure 1, are embedded, respectively, into the Story Manager and into the Stage Manager (where they are stored in the Media Repository). The system loop orchestrates the interaction of the four modules in the following way:

1. The Audience Manager accounts for the reaction of the audience. The system relies on a face expression recognition API.⁵ However, in the two case studies

⁵ <https://azure.microsoft.com/en-us/services/cognitive-services/emotion/>

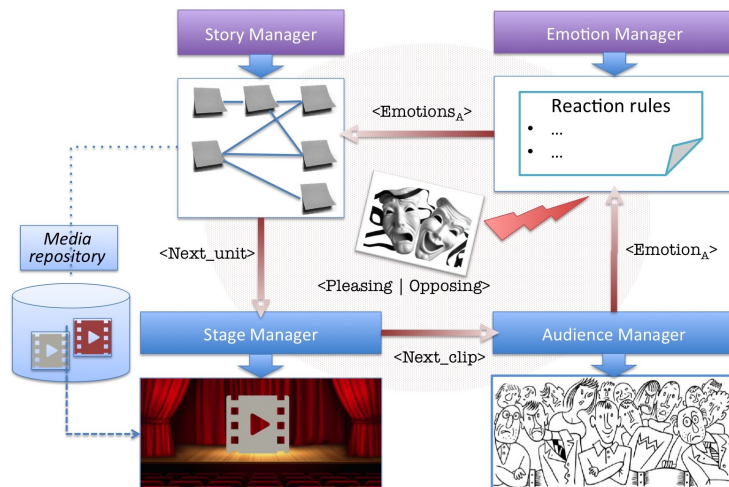


Fig. 4. The architecture of DoPPioGioco. Red lines represent the control flow of the system.

below, we have simulated the audience reaction through a random selection of emotions, displayed as pictures taken from a preloaded set provided by the participants.

2. The Emotion Manager takes as input the audience's response and the attitude selected by the performer for the reaction, and applies the Reaction Rules to compute the candidate emotion family, which becomes the input to the Story Manager.
3. The Story Manager selects the next unit based on the story graph and the emotion family selected by the Emotion Manager.
4. The Stage Manager takes as input the unit selected by the Story Manager and sends it to the player. The Stage Manager provides the properly called intelligent prompt: by consulting the story graph to select the possible continuations, it filters out the units that are not emotionally consistent with the attitude toward the audience's emotional response selected by the performer (and thus, with the emotion family determined by the Reaction rules).

Each module of the system is implemented as a web service (written in PHP and mySql), so as to allow the portability of the system across different devices and media. The current interfaces have been developed as web pages and rely on the Ajax technology to support a fluid interaction with the user during both the real time and the annotation phases.

5 Case Study and Lesson Learned

During the academic years 2015-16 and 2016-17, DoPPioGioco provided the online platform for a writing lab targeted at graduate students in Arts and

Media at the University of Turin. The goal of the lab was to teach the students to design and produce an interactive story by accounting for the emotions in the performer/audience relation. The actual performance has not been staged live, but each student could play the performer role through the Stage interface, being faced with a simulated audience (audience faces generated by the students themselves, 12 pictures, one per emotion in the GEMEP model). The chunks were produced and delivered in a multimedia format, to be self-contained with respect to the narrative content, and not as a simple suggestion to the performer in remembering the chunk addressed.

This lab was a test-bed for assessing the potentials of the platform for training goals. The staff consisted of a writing supervisor, a media producer, and a developer. The writing coordinator was in charge of coordinating the conception of the story, the design of the characters and the editing of the story in textual form. The producer coordinated the production of the audiovisual clips that realize the story units. The developer was engaged in debugging the software system and implementing small design modifications during the conduction of the laboratory. The first edition included 18 students. In this edition the story was created from scratch leveraging on the students' ideas. One main idea was selected through a contest held among the participants. After agreeing on one specific story plot, the class collaboratively designed the story graph. That means that each student was given a branch of story to be developed in a sequence of single story units. Finally, the students were divided into small teams and were assigned the editing, annotation and production of the story units. The final product was delivered online. In this edition of the lab, the story, titled "Clark", was a spooky thriller set in a hospital haunted by a ghostly presence. The main character of the story is a hospital worker, Clark, who comes across some apparently paranormal events. With the help of a mysterious and poised young girl, Magdalene, Clark investigates the origin of the events in a climax of horror. But the tale is nothing but a dream made by Clark before his first day in the new job: the alarm clock rings, and he wakes up. The story could evolve in different narrative ways, according to the audience reactions; however, after running several interactive sessions – where the students alternated between the roles of the performer and the audience, we realized that the story was sometimes perceived as sparse and lacking unity of action.

In the second edition of the lab, with 17 students, we leveraged on the experience of the previous year. First, we decide to use a looser type of narration, because the system is designed to help the performer to manage the emotional response of the audience, rather than guiding the narrative cause-effect chain. Second, we switched the students' focus on the emotions delivered rather than on the creation of the narrative sequence. Thus, we decided to provide a ready-made storyworld in the form of a repository that describes the life in a small American town in the Sixties, by employing the software developed by James Ryan for the performance/installation "Bad News" [16]. Using this software, we produced a specific setting that can be visualized as an exhaustive list of characters, occupations, relations, conflicts, family trees, etc.. for a fictional smalltown.



Fig. 5. Students as authors that physically simulate the graph of story units, also identifying responsibilities for unit–unit connections and improvising the story continuations before actually creating them. Notice the print graph of Figure 3 on the upper left part of the image.

Moreover, yet beforehand, the writing coordinator designed and fixed the story plot structure in a form of a graph (Figure 3) in which every story unit belongs to a specific set of emotions (see GEMEP model above) and linked to the others units accordingly. Once in the lab, we asked the participants to select a genre according the main conflicts listed. The choice was a mix between “romance” and “business competition”. Thus, after a preliminary discussion, we came out with a story, titled “Hot bread”, centered around two bakeries (an old established one and a new rampant one) mixed with a romance between the two main characters. Therefore, we divided the participants into authoring teams, after having experienced the unit connections in a physical performance simulation (see Figure 5). Each group was assigned to write a specific set of units, according the following constrains: 1) being loosely consistent with the antecedent units; 2) being consistent with one of the emotions of the given set. This solution was successful in helping in the process of a collaborative writing (i.e. to produce story unit that were consistent one another) and to devise a multi-plot that was nevertheless aimed toward a closure. Indeed, the characters go through changes in their personal lives as a result of their professional and relational crossroads and the story emerged as inspired to serial formats (such as TV fiction).

The experience with DoPPioGioco was crucial to transmit to the students the notion that writing has to take into account the architecture of the interactive storytelling system (in our case, the suggestions issued by the intelligent prompt). In particular, how the writing methodology can take full advantage of the potential of the platform and how to identify the most suitable narrative

type for it. In general, the students appreciated the approach underlying the system and correctly grasped the relation between the annotation of the units with the emotion tags and the generation of the story in the interplay with the audience, which allow the system to trespass the invisible “fourth wall” posited between performer and audience. The story graph proven to be successful in securing the connection among units and the branching for each type of emotion defined by our set. As a result of the changes in the design and genre of the story, the number of units dropped in the second edition (156 to 105); also, due to the more thorough design of the story, in the second edition the duration of the story was a function of the intensity of the emotions: more intense audience responses (remember from Section 4.3 that intensity was one of the two axes of the emotion system) led to shorter stories, i.e., stories composed of lower number of units. The tension between the creation of well crafted, consistent story line and the emotional response of the audience was both a challenge and an opportunity for the trainees. In particular, the factorization of the emotions in story editing, and the dualism between the roles of story editor and producer, were suitable to promote a change of paradigm in their design and editing practices. The writing labs conducted with the system provided some useful insights for the redesign of the system. Putting forth emotions in the design of the story actually went against the standard practices previously learned by the students, mostly based on causal connections: as a result, designing the story in terms of emotions tended to put at risk the logical consistency of the incidents, requiring a careful control on the story graph. To alleviate this problem, a tool for checking their consistency, similarly to [9], may be added to the system in the future.

In order to train the students to cope with unexpected cases of audience responses, in both editions of the writing lab the recognition of the audience’s emotions was replaced with a simulation of the audience. The random generation of the emotions of the audience, even if sometimes frustrating, turned out to be effective for teaching purposes, since it led the students to face the extreme consequences of the interaction with the audience. However, in order to reduce the gap between the author’s expectations and the reactions of the audience, the set of emotions may be reduced to a more basic set. As future work, we plan to extend the use of the system to other storytelling forms, such as teaching and presentation, and in mixed forms, such as edutainment and infotainment.

6 Conclusion

In this paper, we described a model of performer–audience relationship mediated by an intelligent prompt, in an interactive storytelling setting. The DoPPioGioco model, implemented as a web–based system, provides the author with a story authoring interface, where story chunks are annotated with emotion tags; on stage, the performer receives from the system an intelligent prompt, computed by taking into account the audience’s response and the performer’s attitude towards the audience. The system has been employed in a class of interactive storytelling authoring, with the goal of training students to shift from a narrative

perspective to an emotion-based perspective in interactive story editing. The focus was mostly on writing the story chunks, accounting for the exhaustive combinatorics of the system and simulating the performance through a stage manager provided via web. The experience was very stimulating for the students, with an improvement of the results attained along two editions.

We are planning to realize real live performances in the next editions, employing devices for capturing the facial reactions of the audience and providing a suitable haptic interface to the performer on stage.

Acknowledgments. The authors acknowledge James Ryan for having provided the software for the city simulation.

References

1. Alrutz, M., Listengarten, J., Wood, M.V.D.: *Playing with theory in theatre practice*. Palgrave Macmillan (2011)
2. Aylett, R., Louchart, S.: I contain multitudes: creativity and emergent narrative. In: *Proceedings of the 9th ACM Conference on Creativity & Cognition*. pp. 337–340. ACM (2013)
3. Aylett, R.S., Louchart, S., Dias, J., Paiva, A., Vala, M.: Fearnot!—an experiment in emergent narrative. In: *International Workshop on Intelligent Virtual Agents*. pp. 305–316. Springer (2005)
4. Bänziger, T., Scherer, K.R.: Introducing the geneva multimodal emotion portrayal (gemep) corpus. *Blueprint for affective computing*. pp. 271–294 (2010)
5. Baumer, A., Magerko, B.: Narrative development in improvisational theatre. In: *Int. Conf. on Interactive Digital Storytelling*. pp. 140–151. Springer (2009)
6. Brooker, W.: *The audience studies reader*. Psychology Press (2003)
7. Gratch, J., Marsella, S.: Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In: *Proceedings of the fifth international conference on Autonomous agents*. pp. 278–285. ACM (2001)
8. Koenitz, H., Ferri, G., Haahr, M., Sezen, D., Sezen, T.Í. (eds.): *Interactive Digital Narrative. History, Theory and Practice*. Routledge (2015)
9. Lombardo, V., Damiano, R.: Semantic annotation of narrative media objects. *Multimedia Tools and Applications* 59(2), 407–439 (2012)
10. Louchart, S., Truesdale, J., Suttie, N., Aylett, R.: Emergent narrative: Past, present and future of an interactive storytelling approach. In: *Interactive Digital Narrative. History, Theory and Practice*, pp. 185–199. Routledge (2015)
11. Mateas, M., Stern, A.: *Façade: An experiment in building a fully-realized interactive drama*. In: *Game Developers Conference*. vol. 2 (2003)
12. Plutchik, R.: *Emotion: A psychoevolutionary synthesis*. Harpercollins College Division (1980)
13. Riedl, M.O., Young, R.M.: Narrative planning: balancing plot and character. *Journal of Artificial Intelligence Research* 39(1), 217–268 (2010)
14. Roberts, D.L., Narayanan, H., Isbell, C.L.: Learning to influence emotional responses for interactive storytelling. In: *AAAI Spring Symposium: Intelligent Narrative Technologies II*. pp. 95–102 (2009)
15. Russell, J.A.: Core affect and the psychological construction of emotion. *Psychological review* 110(1), 145 (2003)

16. Samuel, B., Ryan, J., Summerville, A.J., Mateas, M., Wardrip-Fruin, N.: Bad news: An experiment in computationally assisted performance. In: Proceedings of ICIDS 2016. pp. 108–122 (2016)
17. Scherer, K.R.: Appraisal theory. Handbook of cognition and emotion pp. 637–663 (1999)
18. Swartjes, I., Theune, M.: An experiment in improvised interactive drama. In: Intelligent Technologies for Interactive Entertainment, pp. 234–239. Springer (2009)
19. Tanenbaum, J.: How i learned to stop worrying and love the gamer: reframing subversive play in story-based games. In: Proceedings of DiGRA (2013)
20. Weallans, A., Louchart, S., Aylett, R.: Distributed drama management: beyond double appraisal in emergent narrative. Interactive Storytelling pp. 132–143 (2012)
21. Zagalo, N., Göbel, S., Torres, A., Malkewitz, R., Branco, V.: Inscape: Emotion expression and experience in an authoring environment. In: Proc. of TIDSE. pp. 219–230. Springer (2006)