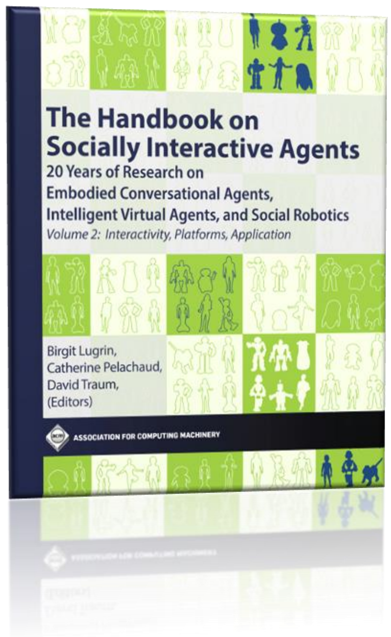
Interactive Narrative and Story-telling

Ruth Aylett



Author note:

This is a preprint. The final article is published in "The Handbook on Socially Interactive Agents" by ACM.

Citation information:

R. Aylett (2022). Interactive Narrative and Story-telling. In B. Lugrin, C. Pelachaud, D. Traum (Eds.), *The Handbook on Socially Interactive Agents – 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics*, Volume 2: Interactivity, Platforms, Application (pp. 463-492). ACM.

DOI of the final chapter: 10.1145/3563659.3563674

DOI of volume 2 of the handbook: 10.1145/3563659

Correspondence concerning this article should be addressed to Ruth Aylett, R.S.Aylett@hw.ac.uk

26 Interactive Narrative and Story-telling

Ruth Aylett

26.1 Motivation

We can look at SIAs in computationally-driven narrative systems from two perspectives. The first is the need for SIAs as characters and/or actors in generated narratives. The second is the need for SIAs themselves to have narrative capabilities.

The creation and telling of stories is thought to be a basic human characteristic. All societies we know of have stories as part of their culture, with dramatic enaction and recitation - often accompanied by music - both pre-dating written forms. The earliest conceptualisation we have in western cultures is that of Aristotle [Lucas 1968] in his *Poetics*, dating back some 2,500 years. The book of this work that survived focused on tragedy (it had a second book on comedy that has been lost) and one of its most important concepts was that of *muthos*. This corresponds to what we might call plot, the structured events and actions that make up a story. Plot has remained a central part of narrative theory, and indeed some might argue that having a plot is what makes a story, a story.

A second of Aristotle's concepts was *mimesis*, which is how the story is represented. The *same* story can be represented in quite different narrative forms - as a song, a play, a computer game, a comic strip. Its presentation could also be organised in different ways, hence ideas such as flash-backs and narrative jumps. Not all plot events necessarily appear in the presentation, they may be implied or referred to. Indeed, in classical Greek tragedy, most dramatic actions happen off-stage.

However, like most distinctions between form and content, there is in practice some blurring. Some narrative forms and some narrative techniques work better with some sorts of story than with others. For example, novels are good at portraying the inner voice of characters, while films are good at portraying their external circumstances. The issues involved in moving a story from presentation as a novel to presentation as a film are frequently non-trivial.

An assumption of Aristotle's, and of much subsequent narrative theory, was that a narrative must be authored. The author was responsible for the content, while they might or might not be wholly responsible for the form. In oral traditions, a narrative might be learned and then presented by other performers. The advent of writing allowed the authoring process to be much more easily disconnected from presentation. In theatrical drama, the actors also carry some responsibility. Thus the concept of *being true* to the author's intent.

Artificial Intelligence (AI) researchers became intrigued by the idea that narrative might be computationally authored more than fifty years ago. Natural language understanding and natural language generation are long-standing fields of investigation in AI, and it was natural for researchers to consider whether a computer could generate say poetry, or a novel, rather than just answer questions about the contents of a database. At this time, computer graphics required expensive specialised equipment, so that initially work was entirely language-based. Just as the researchers of the day investigated natural language via grammars, it seemed feasible that *story grammars* might support the generation of narratives [Rumelhart 1975].

However, the 1977 system TALESPIN [Meehan 1977], one of the earliest story generation systems, implemented a different approach. It used a set of rules and goals to output written stories a few sentences long in the style of Aesop's fables. While the quality of the stories as stories was low, its innovation was to give characters goals that motivated them to act and interact. There was therefore no explicit pre-presentation plot, though [Meehan 1977] makes it clear that goals were chosen with the aim of generating specific stories.

There has been a continuing tension between the idea of computer technology as a presentation mechanism for an authored story, and the much more radical idea of a computer authoring a story itself. How could an SIA be part of a computationally-presented pre-authored narrative? A pre-authored plot necessarily constrains the role an SIA might play. There is limited scope for autonomous action selection, since the plot determines the overall actions of the story at some level of abstraction. However, depending on the form, there may be scope for expanding plot actions at presentation time. Thus actors can in principle determine where exactly they move to on a stage (though this is usually decided in advance by a director), what expressive behaviour they display and how they say their lines. An SIA could therefore have a small set of autonomous capabilities subject to the plot. We could describe this as *weak* SIA involvement.

However, the classical authoring assumption, though dominant, has never been universal, and improvisational narrative runs at least as far back as authored narrative. Consider children's play, or social narratives constructed in turn by successive speakers, or dramatic improvisation, *improv*. In these forms, the story does not have a plot when the presentation starts, though it could be said to have had one by the time the presentation finishes. Plot and presentation proceed hand-in-hand and the narrative is co-authored interactively. If we imagine SIAs as possible co-authors, driven as in TALESPIN by their own goals, then we have what we might call *strong* SIA involvement.

The growth of computer graphics, and in particular the immersive display systems of 1990s Virtual Reality, gave a powerful impetus to computer-driven narrative research. The science fiction television series, StarTrek, became a particular inspiration. It included a rest-and-recreation area for crew on long voyages called the *Holodeck*. Here crew could enter fictional graphical worlds (handily equipped with solid objects and real physics) and interact with intelligent graphical characters in narrative experiences. To many researchers, Virtual Reality seemed like a technological basis for a real Holodeck, forcing the questions of how to create interactive narrative to the forefront [Murray and Murray 2017]. This was when the field really took off. The new technologies seemed to offer a wholly novel genre of narrative experience. We will see below however that the addition of interactivity has also had a substantial impact on the difficulty of the task.

So far we have been motivating the use of SIAs as actors in virtual graphical dramas. These might be for entertainment, or they might have a pedagogical purpose. However the reason that all societies engage in narrative production and performance is a deeper one. Narrative structuring is thought to be a fundamental aspect of human self-definition through what is known as *autobiographical memory* [Neisser 1988]. This is the structuring of personal experience into a story-of-self and a story-to-self.

While it relates to the individual, autobiographical memory is very clearly generated through social experience, and gives us a basic story-telling capability. However, as well as being generated through social interaction, autobiographical memory also supports it. Human-human interaction is full of personal anecdotes, and the telling of such personal narratives is involved in the building of trust and rapport. For these reasons, SIAs in any domain that are expected to function over longer than a half-hour experiment, possibly for months or even years, should be equipped with such a memory and the story-telling capabilities that it supports [Pointeau and Dominey 2017] [Ho and Dautenhahn 2008] [Lim et al. 2009].

26.2 Models and Approaches

The extent to which instances of stories fit theoretical definitions varies. The long history of story-telling includes myths and legends, parables, comedies and satires, tragedies of all sorts, narratives in which one story jumps into another partway through, soap operas, anecdotes, miracle plays, absurdist dramas, comic strips, live-action role plays and a large set of other genres. And that is without including film and musical drama. Expecting one theory to cover all of them over all that time and all those cultures is a big ask.

However, you might think that since narrative has been around for so long, by now there would be an uncontested definition. The most widely accepted definition in western societies would be Aristotle's, focusing on a structured set of events, where 'structured' usually implies some form of causal linking. However, on its own, this would include weather forecasts as narrative as well as most scientific experiments. These are not themselves seen as typical of narrative, though they may be 'turned into' narrative. Thus the story of the discovery of penicillin, or the 1987 UK weather forecast in which the forecaster jocularly dismissed the idea of a hurricane only for one to hit the country a few hours later. These *storified* versions [Aylett 2000] seem to turn a sequence of events into a plot by having them impact characters. If for *characters* we read SIAs, then here is a possible in.

Aristotle did include characters as part of the necessary content of narrative, but in his view they were subsidiary to the plot, forming a mechanism through which the plot was revealed. This was very much in line with Greek tragedy in which characters had limited agency, being subordinate to Fate. Adding a requirement for stronger character agency, and a further requirement for events to impact characters - change their state - as well be caused by them, better matches modern story forms. Attempts to reapply Aristotle to modern genres, whether by media theorists [Laurel 2003] or system builders [Mateas 2000] have focused on redefining the relationship of characters to plot.

The emphasis on structure as a defining element of narrative was taken up by an early 20thC school of narratology known as *structuralism*. Working at a much finer granularity than Aristotle, structuralists looked for elements in common over all narratives. The Russian structuralist Vladimir Propp has especially influenced some in computational narrative systems with his 1928 analysis of Russian folktales [Propp and Scott 1968] in which he proposed some thirty common structures. Each structure bundles characters and events. See for example structure no2: *Interdiction*: A forbidding edict or command is passed upon the hero ('don't go there', 'don't do this'). The hero is warned against some action; or structure no16:*Struggle*: The hero and villain meet and engage in conflict directly, either in battle or some nature of contest. Similar ideas from Joseph Campbell in late 1940s America [Campbell 1956] produced the idea of The Hero's Journey.

Computational modellers are attracted to structures as this is what computers are good at. Propp and Campbell might be very good starting points for a system based on fairy stories [Paiva et al. 2001] or for authoring a quest-based computer game. Not only this, the technology of AI planning acts as a very convenient implementation mechanism, concerned as it is with coherent action sequences meeting specific goals. The UNIVERSE system of the 1980s [Lebowitz 1987] used planning to create *plot outlines* for interpersonal melodramas, aka soap operas. Taking Aristotle's plot/presentation distinction, its outputs were very much at the plot level, since generating character-level actions, in the form of dialogue, was considered a problem to be solved later. As a result they read more like summaries of stories than stories themselves.

It should be no surprise that structuralist approaches are rather genre-specific, and do not generalise well across genres. Propp targeted Russian fairytales, not soap operas. Moreover, as with Aristotlian pre-authoring, they are not easy to reconcile with interactivity. They assume a linear narrative in that the underlying plot forms a single sequence of actions from beginning to end.

Interactivity was both the inspirational contribution of the Holodeck idea, and a major theoretical and practical problem for computational narrative. It presents what has been called *the narrative paradox* [Aylett 2000], a conflict between pre-determined narrative structures – especially plot - and the freedom a graphical world offers a user for spatial navigation and interaction, integral to their feeling of physical presence and immersion. It is worth noting that the superficial visual similarity between interactive graphical worlds and film has led to substantial influences from the latter, especially visible in computer games with their video cut-scenes. This is in spite of film being in general one of the least interactive of forms. Even its actor-level behaviours are under the control of a director through multiple cuts during filming as well as later editing, and the final spectacle is intended for spectators, not participants.

Broadly, the whole field of interactive narrative falls into a plot-dominated camp that attacks the narrative paradox by accommodating limited degrees of freedom for the characters, and a more radical character-based camp that examines how far and under what conditions character interaction can produce spontaneous narrative structure. This has become known as *Emergent Narrative* [Aylett 1999] [Louchart and Aylett 2004] and relies on the idea that just as physical structures can emerge from interaction in the natural world so narrative structure can emerge from human interaction under specific conditions.

In some ways this is yet another rerun of the top-down versus bottom-up arguments that popup in many branches of AI research. The top-down approach guarantees meeting structural goals but at the expense of responsiveness and flexibility, while the bottom-up approach is flexible and responsive but lacks guarantees about outcomes. The bottom-up emergent narrative approach has helped to direct narrative research into new channels however. It focused on improvisational narrative forms that had been relatively neglected: interactive theatre [Boal 1995], theatrical improv [Magerko et al. 2010] and role-playing games [Louchart and Aylett 2003] - table top, live-action [Tychsen et al. 2006], and massive multi-player online role-play games (MMORPGs).

Plot-based approaches can be more or less pre-authored, depending on the level of abstraction of the narrative goals to be met. If one follows Aristotle, then every character action would be specified, with characters at most having scope to interpret the details of execution. On the other hand, a hierarchical AI planning perspective could work at the level of 'boy meets girl' allowing for real-time expansion of exactly how this happens. In both cases the narrative would remain essentially linear.

However the most popular approach in plot-based narrative to supporting a degree of character choice abandons linearity for narrative branching [Riedl and Young 2006], along the lines of the 'choose your own adventure' books of the 1980s and 1990s. These coincided with the arrival of hypertext as a technology. A section of literary theorists became very interested in hypertext, seeing it as a way of authoring interactive narrative [Landow 1991] without the need for advanced technical skills.

As a hidden form of menu selection [Delmas et al. 2007], it is widely used in computer games to provide a degree of replayability, although in most cases branching occurs within well-defined subsections of the game which then all branch back to a main path. The plot can no longer be represented by a chain of actions but becomes a directed graph. This does offer limited autonomy to both user and characters, even if this may boil down to making a choice between a small number of narrative branches.

Character-based approaches are by definition not pre-authored in the same way, though they do require a specified world in some initial state, and motivated characters. They also require mechanisms to shape interactions into something that feels like a coherent story to participants. A story director agent is a typical solution [Weyhrauch 1997]. Such a director may have variable powers: in some systems it might decide a character's action for them at a critical point, or might be confined to acting on the storyworld in the style of an RPG Gamemaster [Tychsen et al. 2009]. A more distributed solution, not incompatible with a director [Weallans et al. 2012], is to equip characters with an understanding of their social role as an actor in a drama. For example, a character could evaluate possible actions for their dramatic impact on others around them [Aylett and Louchart 2008].

Clearly, there is much more scope for SIAs in character-based approaches; indeed the whole approach hinges on SIA capabilities. In its most radical form, its solution to interactivity is to put the participating user on the same footing as any other character, generalising their interactive freedom. This is also a sharp move away from the idea that the narrative is performed for an audience of spectators, an assumption of most narrative theory. In turn this requires new methods of assessing narrative success, since it is perfectly possible for an emergent narrative to look incoherent to a spectator - depending on their point of view - while feeling coherent to a participant [Aylett and Louchart 2007]. Consider the narrative experience of an MMORPG player on a group quest and whether there is in fact any room for a detached audience.

The key difference between interactive narrative and story-telling is that in the former, characters act out the story, while in the latter a single character, the story-teller, presents the story verbally. Interactive narrative supposes the user to be a participant in some form. Story-telling supposes an audience, and the storyteller has explicit control of the usually pre-authored narrative, though they might alter their telling in response to audience reactions or invite specific participative actions [Kistler et al. 2011].

This poses very different requirements for SIAs. For interactive narrative, the focus is on the actions characters can take to forward the narrative; in story-telling an SIA narrator above all needs appropriate expressive behaviour with which to dramatise the telling. Of course, if the text is pre-authored, it can also be annotated with such behaviour, using an appropriate markup language [Niewiadomski et al. 2009].

26.3 History / Overview

Before about 2000, most work in this field focused on story generation and not interactivity. An influential system of this period was Minstrel [Turner 1993], which unlike the story-grammar approach, applied case-based retrieval and adaption. As its name suggests, it focused on stories of knights and chivalry, and used an author-level planner to call a case-based reasoning system. This refined the story-graph. For example it could insert dramatic support for a major event or an introduction for a new character. It was sufficiently comprehensive to be reconstructed by a different research group more than fifteen years later [Tearse et al. 2010].

By 2000, new groups of researchers had been inspired by the idea of building something like the Holodeck [Cavazza et al. 2000] and interactivity became a very prominent theme. A minimum requirement for supporting interactivity was some element of real-time generation with which a user could interact. This did not per se exclude the earlier story-telling approaches or authorial control if they could be run incrementally and flexibly, but gave an impetus to AI planning as a generative mechanism whether from a plot or a character-based perspective [Riedl and Young 2006], [Aylett et al. 2006].

Interactive technology may come in different forms. The Holodeck supposed an immersive graphical environment, and the user acting alongside SIAs on an equal footing in a shared space. Most research groups did not have access to this type of graphical environment and so we will see that many systems were constructed on desktop machines with keyboard interaction, modes that were easy to apply within free games engines. As a halfway house, and especially as interactive devices like the Microsoft Kinect became available, systems used big screen projection and allowed users to interact by gesture. Some systems dispensed with graphic visualisation altogether and used text.

The degree of narrative interaction may also vary. The Holodeck assumed users acted as full characters inside the narrative. We will see that other less interactive roles are also possible as well as being technically more feasible.

A pioneering instance of the character-based approach, FearNot! [Aylett et al. 2005] cast the user as an 'invisible friend' of one of the characters in a desktop virtual drama. This stance was based on the idea of the *spect-actor* from Forum Theatre [Boal 1995]. The application was targeted at educating 9-11 year olds against bullying and there were obvious ethical reasons for not allowing a user to participate directly.

The narrative was episodic and between episodes in which one character was bullied by others, the user could advise their 'friend'. The SIAs were driven by a comprehensive affective architecture [Dias et al. 2014] and action selection depended on their modeled emotional state. The user's advice influenced, but did not determine, what the character would do in the next episode by impacting parameters in its architecture. Advice content also determined the selection of an episode setting and cast in which the advice was likely to apply.

A different approach was taken in another desktop system, I-STORY-TELLING, using a scenario based on the US sitcom 'Friends' [Cavazza et al. 2002]. Here a character was driven by a hierarchical planner, and the user was able to change the state of the world in ways that would impact their plan. Thus if one character was looking for the diary of another, with a view to checking where to take them on a date, moving the position of the diary would impact whether they did or did not meet the diary's owner along the way. This might then precipitate replanning, and a change in the narrative. Here the user is playing the role of a *ghost*: not visible to characters in the story world but able to change its state. A further consequence was that just initialising the system differently could produce different narratives even without user intervention.

The Mission Rehearsal Exercise [Swartout et al. 2006], a prototype training system built for the US military, took a different interactive approach. A single participant played the role of an army lieutenant but interacted with the scene in plausible military fashion via limited speech with a sergeant character. This was a conversational SIA with an affective architecture [Gratch 2001] who would relay commands changing the state of the story-world, along a small set of possible branches. Other characters carried out simple pre-scripted behaviours. The system used projection onto a screen and surround-sound to produce a feeling of immersion, with characters at approximately life size.

A related system by the same group, SASO, [Core et al. 2006] supported direct interaction between the participant and an SIA playing the role of a doctor in Iraq, whose clinic has to be moved to make way for military operations. The participant was able to communicate with the SIA using speech. The aim of this training system was to develop negotiation skills; the narrative was again small-scale and one-episode long. In both these systems, the narrative was dominated by the training objectives, much as it is in human training role-plays.

Similar in inspiration as interaction technologies matured and became more accessible, were the *ORIENT* [Aylett et al. 2009] and *Traveller* [Degens et al. 2013] systems. Both had SIAs driven by the same architecture as in the earlier FearNot! and both were episodic in construction. Like the systems just mentioned, they combined narrative with pedagogy, in this case education in cultural sensitivity. ORIENT combined three participants around the age of 14 into a team by giving them complementary interaction devices: one with a touch-sensitive Dance Mat for navigation, one with a game-based WiiMote for gestures, and one with an RFID-equipped mobile phone to allow objects to be passed from real to virtual world. The SIAs represented frog-like aliens with unfamiliar cultural mores whose world was threatened by an asteroid. The participants were firmly on the real-world side of the screen but could communicate across it, and social interaction between the participants was viewed as a significant component of the pedagogy.

The later Traveller system used simpler but more functional interaction technology in the Microsoft Kinect. A single participant of young adult age was represented by an avatar inside the story-world, with gestures used to select between actions within scenes, giving a branching narrative. A further innovation in interactional technology was produced in a digital version of part of the novel madame Bovary [Cavazza et al. 2007]. Here physiological sensing was used to infer the possible affective state of the user and generate choices in a branching narrative.

Even more ambitious interaction technology was used in the US *Gunslinger* installation [Hartholt et al. 2009], using an augmented reality approach. A real world Wild West saloon set incorporated *flats* - large video panels - showing life-size SIAs. These could respond to gesture and speech, and the participant was equipped with a cowboy hat containing tracking equipment and a holster and six-shooter that worked like those in video arcades. However one participant [Murray 2017] reported that there were not enough narrative and interactional cues.

The very ease of navigation in the real world compared to that in virtual worlds perhaps put even more onus on interaction with the characters. This also illuminates a narrative problem associated with free participant navigation in the storyworld - they may not be where the story is happening when it is happening. Character-based approaches must design participant goals carefully in order to structure their narrative experience. This is very much a lesson from Live Action Role-Play in which briefing participants about their character's back-story and goals is a key part of the setup.

With multiple participants, the distinction between plot and presentation, referred to as Fabula and Sjuzet (or discourse) by structuralists like Propp [Propp and Scott 1968], becomes inadequate. This is because there is no longer a unique plot in a world in which many characters and many events co-exist. Three rather than two levels [Swartjes and Theune 2006] have therefore been suggested.

The Fabula becomes a causal network of all events that took place in the story world from beginning to end. However the plot is a subset of these events from a specific character/participant perspective, and thus no longer unique, with one Fabula supporting many plots, as in MMORGs. The discourse, as before, is how these events are presented. The Virtual Storyteller [Theune et al. 2003] supported this conceptualisation, creating stories with a collection of SIAs. Inspired by Improv, SIAs were given the capacity to alter the storyworld, for example by adding required objects, in order to create a coherent plot for themselves.

The most widely influential system of this period remains the desktop system, Facade [Mateas and Stern 2003]. As a self-contained experience of 15-20 minutes it was distributed and run well outside the usual circles of researchers in the field. It made the participant a neighbour of a married couple invited round for drinks. The narrative is driven by conflict between the two SIAs along the lines of the play *Who's afraid of Virginia Woolf?*, drawing the participant into a set of strong affective conflicts. Unlike the systems discussed so far, the drama was was not driven by autonomous characters. The authors held to the Aristotlian principle of a plot as mainspring of narrative structure [Mateas 2000], in opposition to strong-SIA concepts.

In order to support a range of user actions, Facade was authored as a large database of plan-fragments called *beats* accessed via interaction rules. Rather then a limited number of explicit narrative branches, it attempted to exhaustively cover the narrative space, which required a very long development period of about five years. A participant was able to freely navigate in the space representing the SIA's apartment, and interacted by typed free text.

If the downside of character-based approaches is an incoherent plot, the downside of plot-based approaches is characters acting 'out of character' in order to fulfill the plot's demands, undermining their believability. Facade dealt with this problem by only authoring beats that maintain character believability. A different approach is to treat character believability as a constraint on a planner generating plot [Riedl and Young 2010]. Work on the Intent-based Partial Order Causal Link (IPOCL) planner incorporated reasoning about character intentionality. It divided planned events into *happenings* - accidents, involuntary reactions to stimuli, and forces of nature that do not require character intentionality - and *non-happenings* that do. It both recorded possible intents characters could have, and temporal frames over which they committed to trying to realise them, allowing non-happening events to be motivated by character goals.

26.4 Similarities and Differences between IVAs and SRs

The reader may have noticed that so far little mention has been made of SRs, and that the systems discussed have all involved IVAs. This is because there is much less work involving SRs and narrative, with what there is dominated by story-telling rather than interactive narrative.

SRs share with IVAs the affordances of embodiment including expressive behaviour, and the ability to exhibit agency. That said, there are also substantial differences, making them harder and more time consuming to work with. IVAs can have perfect information about the graphical world, including any other IVAs. This world is in principle a wholly controllable environment. Mobile SRs are physical artefacts in a imperfectly controllable physical world, with fallible sensing and actuation capabilities. A basic requirement in drama of knowing where you are on the stage and what is within your field of view is trivial for an IVA and very tricky for an SR.

Furthermore, a constraint that is seldom mentioned is that mobile SRs have a battery capacity of between 30 minutes and a few hours, depending on how much activity they engage in. After which they typically take a few hours to recharge. Most robots are also much more expensive as single artefacts than graphical characters, are available in limited numbers, and require relatively large spaces in which to operate. Larger robots require accessible engineering support and are expensive and difficult to transport. Touring a production with a cast of multiple robots could not be undertaken lightly. Even SF films have never used autonomous robots as actors.

Historically, robots were very much task-focused, exploiting their basic ability to operate physically on the world. While SR research focuses on human-robot interaction (HRI), most of it also operates within requirements for useful function. This may be therapeutic, educational, or assistive, but outside of specific companies (Disney being the prime example: [Kober et al. 2012]) is less often narrative or dramatic. Existing entertainment examples, as in theme parks, are nearly all animatronic rather than exhibiting any degree of autonomy [Lu 2012].

Many comparative experiments suggest that SRs also have a greater social impact on users than IVAs given they share the same physical space [Wainer et al. 2006]. However the potential for physical impacts makes mixing mobile robots and people in social spaces potentially hazardous. In such shared spaces SRs are normally kept stationary or move very slowly with audible warnings, neither being very compatible with a dramatic suspension of disbelief.

For all these reasons, the use of mobile robots as actors is rare and of autonomous robot actors even rarer. A notable exception was the work of a group at Carnegie-Mellon at the end of the 1990s [Bruce et al. 2000] in which two robots autonomously played out an improv exercise in which one actor wants to leave the room and the other has to find ways of preventing them. By giving the robots different personalities, this produced varying performances. An autonomous robot was also used in a drama involving a single robot and actor in France [Lemaignan et al. 2012], mainly designed to demonstrate to an audience the difference between robot and human perception.

There has been more recent work in Japan, which has a long tradition via Bunraku puppetry of using mechanical devices in drama. However this tradition may predispose work to tele-operation [Chikaraishi et al. 2017]. A 2011 US performance of Shakespeare's Midsummer Night's Dream with seven flying robots acting as alter egos to the fairies in the play also used tele-operation [Murphy et al. 2011]. A further US example involved high-level tele-operation of the robot at the level of action selection rather than direct physical manipulation [Zeglin et al. 2014], though this lower level was available in case it was needed.

Where researchers engage in dramatic work with large mobile robots, the motivation is usually the study of expressive behaviour and more flexible HRI [Knight 2011]. Robot theatre requires dramatic rather than naturalistic behaviour, and this in the context of platforms with much more limited expressive capabilities than those available for an IVA. The study of IVA expressive behaviour is substantially based on human-looking IVAs and, as discussed in a reference to Chapter 8 on "Multimodal Behavior Modeling for Socially Interactive Agents" [Pelachaud et al. 2021] of volume 1 of this handbook [Lugrin et al. 2021], often draws on accounts of facial expression from psychology.

SRs may not have a face at all, or if they do, one that supports restricted or even zero facial animation. Lack of expressive speech - discussed in chapter 6 on "Building and Designing Expressive Speech Synthesis" [Aylett et al. 2021] of volume 1 of this handbook [Lugrin et al. 2021] - is also a major expressive deficiency, and becomes even more of a problem in story-telling systems dependant primarily on speech. These are issues that may well be better addressed by the conventions of animated characters [Johnston and Thomas 1981] [Ribeiro and Paiva 2012] than of human-like behaviour.

There is growing interest in using the development of robot-based drama as an educational process in its own right [Ryokai et al. 2009], [Bravo Sánchez et al. 2017], [Sullivan et al. 2017], [Barnes et al. 2020]. Here, the quality of the dramatic output and the level of autonomy of the robot actors is less important than the classroom-based development process and its pedagogic outputs. This work prioritises the use of small robots already available in education, usually with quite limited capabilities. It may also mix them with other physical elements - tangibles or even cards, supporting narrative development by child users [Fischbach et al. 2018].

Story-telling robots form a separate narrative domain that is in general far less technically fraught. Audiences are small, often children, who are more tolerant of technical failures, and usually involve what are known as *desktop robots*. These are much smaller than robots designed to move around a room and they remain in one place. They do usually have movable heads, are often able to use glance, and sometimes have expressive facial features. They are much less expensive and complex and can use mains electricity. On the other hand story-telling creates a very strong focus on the expressive behaviours of the storyteller, since it is this that animates the story and engages the audience [Conti et al. 2017]. Deficits in expressive speech become especially significant and many systems have used recorded human speech as a result.

In principle, a robot telling a specific story could be wholly pre-programmed, with the text annotated with suitable head movements, glance and appropriate facial expressions. However human story-telling is a social activity in which the storyteller takes account of audience reactions to modify their performance and possibly how the same story is told. Given the current limitations on processing user reactions, this is still very challenging, especially inferring an audience's emotional reactions [Costa et al. 2018].

A constructive approach to education sees story-telling by pupils as a source of active learning [Catala et al. 2017a]. This means that story-telling robots need not themselves tell the stories and some of the technical issues can be mitigated by social interaction between the human participants. An SR can provide an inspiration or prop for child story-telling [Plaisant et al. 2000] and the feasibility of co-operative story-telling is also currently investigated [Sun et al. 2017].

26.5 Current Challenges

Most researchers in the field would agree that the early excitement about the possibilities of the Holodeck has not resulted in many fully realised systems that could be enjoyed for themselves as narrative experiences rather than as demonstrations of research ideas [Murray 2017]. Facade, as discussed above, remains a rather isolated example, and is known to have taken five years to develop. FearNot! with 44 possible episodes, was developed and evaluated over the course of two three-year projects.

This is in spite of a great deal of work on authoring systems. The effort involved in developing an initial digital narrative system has led to most groups working in the area over a period of time to develop their own tools to make this easier. Generalising a specific system into a more generic architecture is an obvious move. Hence FAtiMA [Dias et al. 2014] from the FearNot! work, and Thespian [Si et al. 2005] and the Virtual Human Toolkit [Gratch et al. 2013] from systems such as MRS, SAPO and Gunslinger. The experience of building Facade lies behind the toolkit Comme Il Faut [McCoy et al. 2010], the extensive base of social behaviour rules used to build PromWeek [McCoy et al. 2012], a narrative game in the style of those that appear in The Sims.

Some work has started at the other end of the process, with an explicitly designed authoring system that incorporates runtime support for its output. However, just as architectures abstracted from specific systems embody the narrative approach and assumptions of that system, so authoring systems also embody specific assumptions and representations. This has always been clear in the authoring systems associated with games engines. Thus the Dungeons and Dragons-based Never Winter Nights of the late 1990s had an authoring system based around locations, navigation between them, objects found in them, and creatures encountered there, either for conversations or battles. On the other hand the first-person shooter engine, Unreal Tournament, of the same period, had a system for authoring shooter episodes.

IDtension [Szilas 2003] is a longstanding narrative authoring system, text-based in its original form. It is plot rather than character based and is able either to generate a story itself or to produce one in real-time with a participant making action choices alternately with the system. Its approach is to present every possible action in its model at every turn, making a search through its knowledge base of goals, actions and obstacles. This can result in more than a hundred possible choices during the story. It was made available to non-experts to produce a complete sample drama, The Mutiny [Szilas 2008]. Other authoring systems, such as STORYTEC [Göbel et al. 2008] may have narrative elements but are primarily game authoring tools, usually for adventure-type games in which a narrative is really a backstory for puzzle-solving.

While invaluable in increasing the productivity of research groups, authoring systems have not so far led to more complete narrative experiences. One obvious reason is the theoretical difficulties we have discussed of reconciling structure and interactivity are still not wholly solved. In computer games (see chapter 27), this issue is generally addressed by restricting narrative interactivity in favour of much simpler forms such as shooting. Narrative branching is still the standard approach in commercial games, along with the use of non-interactive video cut-scenes. Non-player characters rarely have the necessary capabilities for autonomous action, in most cases not even a memory of past interactions. However, two other reasons are at least as important. The first is the need for content, and the second the technical skills required to develop such systems.

In the other graphical genres - film and computer games - large numbers of people are involved in generating content. Indeed, many games companies do not distinguish between pre-authored narrative and pre-authored graphical content. They are as wary of open-ended narrative as they are of machine learning, and for similar commercial risk-avoidance reasons. Researchers, on the other hand, seldom have access to large-scale graphical content and even more seldom have such creative skills themselves. Routine content production is not a publishable research task. This is one reason for widespread use of game engines, since they come with free content or a user base that generates and shares such content.

In general, the greater the degree of user interactivity, the worse the graphical content problem becomes. If users can change the state of the story world then new states must be visualised for them. If story characters can also change the state of the world then the problem multiples. Moreover, the greater the roles to be played by characters, the greater the repertoire of actions they require, all of which must also be visualised. Not only this, characters nearly always require natural language interaction capabilities, and this requires actual language content, whether pre-authored phrases, lexicon and grammar, or a machine learning-based dialogue system. Character-based narrative approaches have more substantial content demands than plot-based approaches, because their aim is to produce a wider range of narrative experiences.

Table-top role-play solves this problem by supplying gamemasters with volumes of background world material, including the milieu, some history and a set of back-stories. This works because there are plenty of gamemasters and roleplayers and using the materials only requires the cost of the book and literacy. Moreover the materials are interpreted by human intelligence. But consider turning the whole of a roleplay universe into shareable graphic materials. This would be a much larger scale enterprise than the materials associated with a game engine and would have to find a common representation - graphical and probably ontological too - that could be used by multiple game engines. Specific MMORPGs do solve the problem - and some, such as World of warcraft, are effectively graphically versions of earlier written source materials. However, these are closed proprietary systems. Open systems like Minecraft allow users to generate graphical materials but do not embody any narrative framework.

The second barrier to fully-realised systems is the level and range of technical expertise required to author them. The large teams involved in film and computer-game production support multiple skill specialisations. Interactive narrative shares similar integration problems, but research teams are rarely large enough to include a full range. While a novel can be attempted by any literate person on their own, the need for an understanding of narrative technology is a barrier to creative non-specialists. Widely usable authoring systems remain a challenge.

The technical experts, on the other hand, are rarely talented constructors of compelling narrative and characterisation. Game engines provide 3D graphical environments with backing code banks, but they are focused on supporting specific game genres and require programming effort to support functionality beyond that. The most-used engines are proprietary and have substantial licencing charges if used for publicly distributed work. The research-based authoring systems and architectures discussed above all require a good understanding of the technology of the system being authored to be successfully used.

The content challenge is currently being addressed in three ways: by procedural generation; by crowd sourcing; and by the application of machine learning approaches. Interactive narratives require a broad range of different types of content, some graphical ('physical' environment, graphical models of objects, character bodies and animations) and some related to the technological approach of the narrative being constructed. For example, a plot-based approach may require content for the branching graph representing the plot; an emergent narrative may require character actions.

Procedural generation is widely explored in computer games research [Hendrikx et al. 2013]. Methods are available for terrain, vegetation, bodies of water, road networks, and urban environments, as well as lower-level textures and materials. Procedural animation systems have been developed for the computer games market and researchers have tried to extend these into forms more suitable for autonomous characters [Horswill 2009]. While this is all useful, it is not narrative, and works better for games in which simple interactions like navigation and fighting are important. Moreover most of these systems are developed as middleware for the computer games industry, often at premium prices, and are not necessarily accessible to researchers. Generation of puzzles for RPG games [Fernández-Vara and Thomson 2012] is closer to narrative but a puzzle is not itself a plot.

AI planning technology can be thought of as a procedural generation system working in real-time [Riedl and Young 2010]. However it too has content: binding symbolic representations of world state to goal and action definitions. While a graphical procedural generation system can produce a building, an ontology is required to label it 'a police station', and a planning system would need an action 'report to the police' as well as a motivating goal to make use of it in a narrative. The planner also needs to know what the consequences of executing the action would be in terms of changes in the state of the world. This is true whether the plan operates at the level of a plot, or is attached to an autonomous character.

The AI Planning research community has worked on the problem of automating the construction of planning knowledge for many years [Cresswell et al. 2013], [Jilani et al. 2014]. However there is only a small overlap between the two research communities, and how to generate narratively-useful knowledge for an AI planner remains a challenge [Porteous and Cavazza 2009].

Crowd sourcing forms a second possible source of content. One might argue that the narrative content of MMORPGs is a type of crowd sourcing, though it formally depends on *quests*, a set of authored goals and constraints. An early initiative leveraging large numbers of participants via the internet [Orkin and Roy 2007] ran more than 5000 online sessions of *The Restaurant Game*. The game allowed participants to play either a customer or a waitress, so that data was collected from human role-play. This data was used to build a narrative representation called a *Plan Network*, a statistical model that encoded context-sensitive expected patterns of behavior and language, with dependencies on social roles and object affordances.

This was intended to provide what it called *common ground*, the world knowledge that could be used in both plot and character based approaches about what actions people were likely to perform with which objects in specific scenarios. The approach had clear parallels with corpora-building in the natural language dialogue community, one of which involved ordering meals online from restaurants.

Other work has included crowd sourcing to directly form a plot-graph [Li and Riedl 2015] and the application of structural ideas from Theatre, in the form of Stanislavsky's Active Analysis rehearsal technique [Feng et al. 2016]. Online rehearsal was also the approach taken by a crowd-sourcing system aimed at collecting character material, including change in emotional state, for an emergent narrative [Kriegel and Aylett 2008].

26.6 Future Directions

Crowd-sourcing content pre-dated the enthusiasm for machine learning which has swept the AI community at the time of writing, and they can be thought of as comparatively small-scale examples of these data-driven approaches. Reinforcement learning, in which actions are tried and rewarded or punished according to their success, is the technique of choice for learning sequences of actions. Deep reinforcement learning (RL), using the current artificial neural net technologies, has been applied successfully to learning a better-than-human performance set of actions in a number of Atari games [Hessel et al. 2018].

Most games genres are however much simpler than narratives, with a small set of actions, a repeating set of contexts, and simple metrics for rewards and punishments. Even so, deep RL requires very substantial amounts of computing power. Recent work required an average of 80 million frames on games running at 60 frames a second - equivalent to about 38 hours of human play on a game a human can learn in tens of minutes. With a wider set of possible actions and contexts that repeat much less often, as well as metrics of 'success' that are very hard to formalise, applying deep RL to generating action sets for characters, or for a narrator-level plot, in graphical environments would be a formidable task. Not least a chicken-and-egg problem of not having enough running narratives in the first place to form a corpus. However if the task is reformulated into a purely language-based one, it becomes somewhat more tractable. As already mentioned, the field of natural language has been using corpora for a long time, and there already exist resources that are useful for learning narrative sequences. One example is a database of film summaries [Bamman et al. 2013]. Natural language corpora can be processed into simpler event-like structures with subject-verb-object plus indirect objects where needed.

A problem of standard recurrent neural network approaches is that they have limited ability to retain a memory of past decisions, so that learning plots results in stories that wander, without much narrative coherence. Recent work [Tambwekar et al. 2019] tries to deal with this by shaping learning towards a final goal, a specified event. As in UNIVERSE and other older systems, this generates a plot description, the basic set of actions in the story. It is some way from the presentation of the story and so far work has not examined how the output could used for an interactive experience.

As discussed above, every research group working in a specific software field over a long period develops architectures (see chapter 16 on "The Fabric of Socially Interactive Agents: Multimodal Interaction Architectures" [Kopp and Hassan 2022] of this volume of the handbook) and tools (see chapter 20 on "Platforms and Tools for SIA Research and Development" [Hartholt and Mozgai 2022] of this volume of the handbook) to reduce the amount of reworking they must do each time they produce a system. In the computer games industry this has led to an associated middleware industry supplying common tools - very expensive ones in general - to many games companies. In a research field like interactive narrative the establishment of common representational standards would help this to happen.

The behaviour markup language standard developed for driving expressive graphical characters [Kopp et al. 2006] is a help. However, its general purpose nature, while a strength, makes it a limited component of an interactive narrative system. A set of narrative ontologies for different styles of interactive narrative might deliver more, but the theoretical divergences in the field make these hard to develop. AI Planning has benefited substantially from a Plan Domain Description Language (PDDL) [Fox and Long 2003] but this appears to have made very little impact on work in digital narrative in spite of the central role played by planners. The development of agreed standards is work that still needs to be done.

As a domain in which many technologies must be integrated, interactive narrative also stands to benefit from developments in many related fields. We have already mentioned the need for expressive speech. The rapid improvement in dialogue systems should also be highly significant (See chapter 15 on "Socially Interactive Agent Dialogue" [Traum 2022] of this volume of the handbook), as also developments in multimodal interaction.

We saw above that the growth in interaction technologies has led to systems with a more ambitious approach to interaction modalities than the keyboard. Multimodal interaction has so far been mostly applied to deliberate participant actions, allowing the use of explicit gesture [Degens et al. 2013] - see chapter 7 on "Gesture Generation" [Saund and Marsella 2021] of volume 1 of this handbook [Lugrin et al. 2021]. Much more could be done to incorporate participants into a narrative experience using facial expression recognition to infer responses, or action recognition and goal inference, especially in augmented reality settings. The use of Theory of Mind processing (see chapter 9 on "Theory of Mind and Joint Attention" [Perez-Osorio et al. 2021] of volume 1 of this handbook [Lugrin et al. 2021]), especially in its simulation variant [Dias et al. 2013], also gives scope for enriched interaction.

While the Holodeck was an inspiration to the field, it also prioritised a specific view of an interactive narrative as a self-contained virtual experience. Rather than requiring participants to operate in the virtual world, one can conceive of computer-supported role-play in which SIAs contribute to a real world narrative experience. This augmented reality view would include the use of tangibles [Catala et al. 2017b] and collective human role-play. An inspiration here would be the Story Room [Alborzi et al. 2000], originally conceived as a space with resources supporting child story construction. Robust technology that could support Live Action Role-Play would also be an interesting way of exploring this idea.

Finally, longlived SIAs, interacting over weeks, months, or even years (see chapter 19 on "Long-Term Interaction with Relational SIAs" [Kory-Westlund et al. 2022] of this volume of this handbook), will absolutely require narrative capabilities. These will organise their interaction memories, as well as creating more pleasant and varied interaction [Bickmore et al. 2009] [Cordar et al. 2014]. Interactive narrative will appear at some point wherever longlived SIAs appear.

26.7 Summary

We have seen in this chapter that narrative and story-telling are fundamental human capabilities. As such, they have formed one of the challenges for AI researchers tackling a wide range of human skills.

We have considered how this led to the field of interactive narrative, and the inspiring example of the Holodeck, focusing on the role of SIAs as characters. We have seen that the two and a half thousand years of narrative theory in the west has not always helped all that much in this field due to the nature of interactivity. We have looked at both plot-based and character-based systems and how they have tried to meet the challenge of the narrative paradox, of reconciling user freedom with narrative structure.

Few of the systems we considered have been aimed at entertainment, with education and training often the driving application domains. Computer games have however had a substantial impact, both through the widespread use of game engines for development, and as a possible route to the uptake of narrative technology. As with other branches of AI this last has been a slow process, with issues relating to creative control and technological robustness acting as powerful braking processes.

The differences and similarities between SRs and IVAs have been discussed, and the additional difficulties and opportunities that working with robots provides. We also looked briefly at the differences between interactive narrative and story-telling and why the latter is more often targeted than the former by SRs. Content creation was identified as a substantial bottleneck in the creation of virtual narratives, and we looked at the ways researchers have tried to meet this challenge.

Finally we have examined how far machine learning may contribute, and whether narrative in augmented reality environments can provide some new opportunities.

Bibliography

H. Alborzi, A. Druin, J. Montemayor, M. Platner, J. Porteous, L. Sherman, A. Boltman, G. Taxén, J. Best, J. Hammer, et al. 2000. Designing storyrooms: interactive storytelling spaces for children. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, pp. 95–104.

M. P. Aylett, L. Clark, B. R. Cowan, and I. Torre. 2021. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition*, pp. 173–211. ACM Press, 1.

R. Aylett. 1999. Narrative in virtual environments-towards emergent narrative. *AAAI Fall Symposium - AAAI Technical Report FS-99-01*.

R. Aylett. 2000. Emergent narrative, social immersion and "storification". In *Proceedings of the 1st International Workshop on Narrative and Interactive Learning Environments*, pp. 35–44.

R. Aylett and S. Louchart. 2007. Being there: Participants and spectators in interactive narrative. In *International Conference on Virtual Storytelling*, pp. 117–128. Springer.

R. Aylett and S. Louchart. 2008. If i were you: double appraisal in affective agents. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3*, pp. 1233–1236.

R. Aylett, J. Dias, and A. Paiva. 2006. An affectively driven planner for synthetic characters. In *Icaps*, pp. 2–10.

R. Aylett, M. Kriegel, M. Y. Lim, J. Dias, K. Leichtenstern, W. C. Ho, and P. Rizzo. 2009. Orient: interactive agents for stage-based role-play. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pp. 1371–1372.

R. S. Aylett, S. Louchart, J. Dias, A. Paiva, and M. Vala. 2005. Fearnot!–an experiment in emergent narrative. In *International Workshop on Intelligent Virtual Agents*, pp. 305–316. Springer.

D. Bamman, B. O'Connor, and N. A. Smith. 2013. Learning latent personas of film characters. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 352–361.

J. Barnes, S. M. FakhrHosseini, E. Vasey, C. H. Park, and M. Jeon. 2020. Child-robot theater: Engaging elementary students in informal steam education using robots. *IEEE Pervasive Computing*, 19(1): 22–31.

T. Bickmore, D. Schulman, and L. Yin. 2009. Engagement vs. deceit: Virtual humans with human autobiographies. In *International Workshop on Intelligent Virtual Agents*, pp. 6–19. Springer.

A. Boal. 1995. Forum theatre. *Acting (re) considered*, pp. 251–261.

F. Á. Bravo Sánchez, A. M. González Correal, and E. G. Guerrero. 2017. Interactive drama with robots for teaching non-technical subjects. *Journal of Human-Robot Interaction*, 6(2): 48–69.

A. Bruce, J. Knight, S. Listopad, B. Magerko, and I. R. Nourbakhsh. 2000. Robot improv: Using drama to create believable agents. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, volume 4, pp. 4002–4008. IEEE.

J. Campbell. 1956. The hero with a thousand faces (cleveland. *World*, 53.

A. Catala, M. Theune, H. Gijlers, and D. Heylen. 2017a. Storytelling as a creative activity in the classroom. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*, pp. 237–242.

A. Catala, M. Theune, C. Sylla, and P. Ribeiro. 2017b. Bringing together interactive digital storytelling with tangible interaction: Challenges and opportunities. In *10th International Conference on Interactive Digital Storytelling, ICIDS 2017*, pp. 395–398. Springer.

M. Cavazza, R. Aylett, K. Dautenhahn, C. Fencott, and F. Charles. 2000. Interactive storytelling in virtual environments: Building the "holodeck". In *Proceedings of VSMM*, pp. 678–687. Citeseer.

M. Cavazza, F. Charles, and S. J. Mead. 2002. Character-based interactive storytelling. *IEEE Intelligent systems*, 17(4): 17–24.

M. Cavazza, J.-L. Lugrin, D. Pizzi, and F. Charles. 2007. Madame bovary on the holodeck: immersive interactive storytelling. In *Proceedings of the 15th ACM international conference on Multimedia*, pp. 651–660.

T. Chikaraishi, Y. Yoshikawa, K. Ogawa, O. Hirata, and H. Ishiguro. 2017. Creation and staging of android theatre "sayonara" towards developing highly human-like robots. *Future Internet*, 9(4): 75.

D. Conti, A. Di Nuovo, C. Cirasa, and S. Di Nuovo. 2017. A comparison of kindergarten storytelling by human and humanoid robot with different social behavior. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 97–98.

A. Cordar, M. Borish, A. Foster, and B. Lok. 2014. Building virtual humans with back stories: training interpersonal communication skills in medical students. In *International Conference on Intelligent Virtual Agents*, pp. 144–153. Springer.

M. Core, D. Traum, H. C. Lane, W. Swartout, J. Gratch, M. Van Lent, and S. Marsella. 2006. Teaching negotiation skills through practice and reflection with virtual humans. *Simulation*, 82(11): 685–701.

S. Costa, A. Brunete, B.-C. Bae, and N. Mavridis. 2018. Emotional storytelling using virtual and robotic agents. *International Journal of Humanoid Robotics*, 15(03): 1850006.

S. N. Cresswell, T. L. McCluskey, and M. M. West. 2013. Acquiring planning domain models using locm. *The Knowledge Engineering Review*, 28(2): 195–213.

N. Degens, G. J. Hofstede, S. Mascarenhas, A. Silva, A. Paiva, F. Kistler, E. André, A. Swiderska, E. Krumhuber, A. Kappas, et al. 2013. Traveller–intercultural training with intelligent agents for young adults. In *Proceedings of the Workshop on Intelligent Digital Games for Empowerment and Inclusion at the Foundations of Digital Games Conference (FDG)*, volume 15, pp. 1–8. SASDG Digital Library.

G. Delmas, R. Champagnat, and M. Augeraud. 2007. Bringing interactivity into campbell's hero's journey. In *International Conference on Virtual Storytelling*, pp. 187–195. Springer.

J. Dias, R. Aylett, A. Paiva, and H. Reis. 2013. The great deceivers: Virtual agents and believable lies. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 35.

J. Dias, S. Mascarenhas, and A. Paiva. 2014. Fatima modular: Towards an agent architecture with a generic appraisal framework. In *Emotion modeling*, pp. 44–56. Springer.

D. Feng, E. Carstensdottir, S. M. Carnicke, M. S. El-Nasr, and S. Marsella. 2016. An active analysis and crowd sourced approach to social training. In *International Conference on Interactive Digital Storytelling*, pp. 156–167. Springer.

C. Fernández-Vara and A. Thomson. 2012. Procedural generation of narrative puzzles in adventure games: The puzzle-dice system. In *Proceedings of the The third workshop on Procedural Content Generation in Games*, pp. 1–6.

M. Fischbach, J.-L. Lugrin, M. Brandt, M. E. Latoschik, C. Zimmerer, and B. Lugrin. 2018. Follow the white robot-a role-playing game with a robot game master. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 1812–1814. International Foundation for Autonomous Agents and Multiagent Systems.

M. Fox and D. Long. 2003. Pddl2. 1: An extension to pddl for expressing temporal planning domains. *Journal of artificial intelligence research*, 20: 61–124.

S. Göbel, L. Salvatore, and R. Konrad. 2008. Storytec: A digital storytelling platform for the authoring and experiencing of interactive and non-linear stories. In *2008 International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution*, pp. 103–110. Ieee.

J. Gratch. 2001. Towards the holodeck: building emotional virtual humans for training. In *Proceedings Computer Animation 2001. Fourteenth Conference on Computer Animation (Cat. No. 01TH8596)*, p. 104. IEEE.

J. Gratch, A. Hartholt, M. Dehghani, and S. Marsella. 2013. Virtual humans: a new toolkit for cognitive science research. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 35.

A. Hartholt and S. Mozgai. 2022. *Platforms and Tools for SIA Research and Development*, pp. 261–304. The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application. ACM, 2. DOI: https://doi.org/10.1145/3563659.3563668.

A. Hartholt, J. Gratch, L. Weiss, et al. 2009. At the virtual frontier: Introducing gunslinger, a multi-character, mixed-reality, story-driven experience. In *International Workshop on Intelligent Virtual Agents*, pp. 500–501. Springer.

M. Hendrikx, S. Meijer, J. Van Der Velden, and A. Iosup. 2013. Procedural content generation for games: A survey. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 9(1): 1–22.

M. Hessel, J. Modayil, H. Van Hasselt, T. Schaul, G. Ostrovski, W. Dabney, D. Horgan, B. Piot, M. Azar, and D. Silver. 2018. Rainbow: Combining improvements in deep reinforcement learning. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

W. C. Ho and K. Dautenhahn. 2008. Towards a narrative mind: The creation of coherent life stories for believable virtual agents. In *International Workshop on Intelligent Virtual Agents*, pp. 59–72. Springer.

I. D. Horswill. 2009. Lightweight procedural animation with believable physical interactions. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(1): 39–49.

R. Jilani, A. Crampton, D. E. Kitchin, and M. Vallati. 2014. Automated knowledge engineering tools in planning: state-of-the-art and future challenges.

O. Johnston and F. Thomas. 1981. *The illusion of life: Disney animation*. Disney Editions New York.

F. Kistler, D. Sollfrank, N. Bee, and E. André. 2011. Full body gestures enhancing a game book for interactive story telling. In *International Conference on Interactive Digital Storytelling*, pp. 207–218. Springer.

H. Knight. 2011. Eight lessons learned about non-verbal interactions through robot theater. In *International Conference on Social Robotics*, pp. 42–51. Springer.

J. Kober, M. Glisson, and M. Mistry. 2012. Playing catch and juggling with a humanoid robot. In *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)*, pp. 875–881. IEEE.

S. Kopp and T. Hassan. 2022. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application*, pp. 77–111. ACM Press, 2.

S. Kopp, B. Krenn, S. Marsella, A. N. Marshall, C. Pelachaud, H. Pirker, K. R. Thórisson, and H. Vilhjálmsson. 2006. Towards a common framework for multimodal generation: The behavior markup language. In *International workshop on intelligent virtual agents*, pp. 205–217. Springer.

J. Kory-Westlund, H. W. Park, I. Grover, and C. Breazeal. 2022. *Long-Term Interaction with Relational SIAs*, pp. 195–260. The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application. ACM, 2. DOI: https://doi.org/10.1145/3563659.3563667.

M. Kriegel and R. Aylett. 2008. Emergent narrative as a novel framework for massively collaborative authoring. In *International Workshop on Intelligent Virtual Agents*, pp. 73–80. Springer.

G. P. Landow. 1991. *HyperText: the convergence of contemporary critical theory and technology (parallax: re-visions of culture and society series)*. Johns Hopkins University Press.

B. Laurel. 2003. The six elements and the causal relations among them. *The New Media Reader*, pp. 49–65.

M. Lebowitz. 1987. Planning stories. In *Proceedings of the 9th annual conference of the cognitive science society*, pp. 234–242.

S. Lemaignan, M. Gharbi, J. Mainprice, M. Herrb, and R. Alami. 2012. Roboscopie: A theatre performance for a human and a robot [video abstract]. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 427–427. IEEE.

B. Li and M. Riedl. 2015. Scheherazade: Crowd-powered interactive narrative generation. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

M. Y. Lim, R. Aylett, W. C. Ho, S. Enz, and P. Vargas. 2009. A socially-aware memory for companion agents. In *International Workshop on Intelligent Virtual Agents*, pp. 20–26. Springer.

S. Louchart and R. Aylett. 2003. Solving the narrative paradox in ves–lessons from rpgs. In *International workshop on intelligent virtual agents*, pp. 244–248. Springer.

S. Louchart and R. Aylett. 2004. Narrative theory and emergent interactive narrative. *International Journal of Continuing Engineering Education and Life Long Learning*, 14(6): 506–518.

D. V. Lu. 2012. Ontology of robot theatre. In *Proceedings of the workshop Robotics and Performing Arts: Reciprocal Influences, ICRA*, volume 2012.

D. W. Lucas. 1968. Aristotle poetics.

B. Lugrin, C. Pelachaud, and D. Traum. 2021. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition*. ACM Press. DOI: https://doi.org/10.1145/3477322.

B. Magerko, C. Fiesler, A. Baumer, and D. Fuller. 2010. Bottoms up: improvisational micro-agents. In *Proceedings of the Intelligent Narrative Technologies III Workshop*, pp. 1–8.

M. Mateas. 2000. A neo-aristotelian theory of interactive drama. In *Working notes of the AI and Interactive Entertainment Symposium*.

M. Mateas and A. Stern. 2003. Façade: An experiment in building a fully-realized interactive drama. In *Game developers conference*, volume 2, pp. 4–8.

J. McCoy, M. Treanor, B. Samuel, B. Tearse, M. Mateas, and N. Wardrip-Fruin. 2010. Authoring game-based interactive narrative using social games and comme il faut. In *Proceedings of the 4th International Conference & Festival of the Electronic Literature Organization: Archive & Innovate*, pp. 1–8. Citeseer.

J. McCoy, M. Treanor, B. Samuel, A. A. Reed, N. Wardrip-Fruin, and M. Mateas. 2012. Prom week. In *Proceedings of the International Conference on the Foundations of Digital Games*, pp. 235–237.

J. R. Meehan. 1977. Tale-spin, an interactive program that writes stories. In *Ijcai*, volume 77, pp. 91–98.

R. Murphy, D. Shell, A. Guerin, B. Duncan, B. Fine, K. Pratt, and T. Zourntos. 2011. A midsummer night's dream (with flying robots). *Autonomous Robots*, 30(2): 143–156.

J. H. Murray. 2017. How close are we to the holodeck? In *Clash of Realities 2015/16: On the Art, Technology and Theory of Digital Games. Proceedings of the 6th and 7th Conference*, volume 5, p. 129. transcript Verlag.

J. H. Murray and J. H. Murray. 2017. *Hamlet on the holodeck: The future of narrative in cyberspace*. MIT press.

U. Neisser. 1988. Five kinds of self-knowledge. *Philosophical psychology*, 1(1): 35–59.

R. Niewiadomski, E. Bevacqua, M. Mancini, and C. Pelachaud. 2009. Greta: an interactive expressive eca system. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pp. 1399–1400.

J. Orkin and D. Roy. 2007. The restaurant game: Learning social behavior and language from thousands of players online. *Journal of Game Development*, 3(1): 39–60.

A. Paiva, I. Machado, and R. Prada. 2001. Heroes, villians, magicians,... dramatis personae in a virtual story creation environment. In *Proceedings of the 6th international conference on Intelligent user interfaces*, pp. 129–136.

C. Pelachaud, C. Busso, and D. Heylen. 2021. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition*, pp. 259–310. ACM Press, 1.

J. Perez-Osorio, E. Wiese, and A. Wykowska. 2021. *Theory of Mind and Joint Attention*, pp. 313–347. The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition. ACM, 1. DOI: https://doi.org/10.1145/3477322.3477332.

C. Plaisant, A. Druin, C. Lathan, K. Dakhane, K. Edwards, J. M. Vice, and J. Montemayor. 2000. A storytelling robot for pediatric rehabilitation. In *Proceedings of the fourth international ACM* *conference on Assistive technologies*, pp. 50–55.

G. Pointeau and P. F. Dominey. 2017. The role of autobiographical memory in the development of a robot self. *Frontiers in neurorobotics*, 11: 27.

J. Porteous and M. Cavazza. 2009. Controlling narrative generation with planning trajectories: the role of constraints. In *Joint International Conference on Interactive Digital Storytelling*, pp. 234–245. Springer.

V. Propp and L. Scott. 1968. Morphology of the folktale. *Austin. University of Texas Pres*.

T. Ribeiro and A. Paiva. 2012. The illusion of robotic life: principles and practices of animation for robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pp. 383–390.

M. O. Riedl and R. M. Young. 2006. From linear story generation to branching story graphs. *IEEE Computer Graphics and Applications*, 26(3): 23–31.

M. O. Riedl and R. M. Young. 2010. Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39: 217–268.

D. E. Rumelhart. 1975. Notes on a schema for stories. In *Representation and understanding*, pp. 211–236. Elsevier.

K. Ryokai, M. J. Lee, and J. M. Breitbart. 2009. Children's storytelling and programming with robotic characters. In *Proceedings of the seventh ACM conference on Creativity and cognition*, pp. 19–28.

C. Saund and S. Marsella. 2021. *Gesture Generation*, pp. 213–258. The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition. ACM, 1. DOI: http://dx.doi.org/10.1145/3477322.3477330.

M. Si, S. C. Marsella, and D. V. Pynadath. 2005. Thespian: Using multi-agent fitting to craft interactive drama. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, pp. 21–28.

A. Sullivan, A. Strawhacker, and M. U. Bers. 2017. Dancing, drawing, and dramatic robots: Integrating robotics and the arts to teach foundational steam concepts to young children. In *Robotics in STEM education*, pp. 231–260. Springer.

M. Sun, I. Leite, J. F. Lehman, and B. Li. 2017. Collaborative storytelling between robot and child: A feasibility study. In *Proceedings of the 2017 Conference on Interaction Design and Children*, pp. 205–214.

I. Swartjes and M. Theune. 2006. A fabula model for emergent narrative. In *International Conference on Technologies for Interactive Digital Storytelling and Entertainment*, pp. 49–60. Springer.

W. Swartout, R. Hill, J. Gratch, W. L. Johnson, C. Kyriakakis, C. LaBore, R. Lindheim, S. Marsella, D. Miraglia, and B. Moore. 2006. Toward the holodeck: Integrating graphics, sound, character and story. Technical report, UNIVERSITY OF SOUTHERN CALIFORNIA MARINA DEL REY CA INST FOR CREATIVE ....

N. Szilas. 2003. Idtension: a narrative engine for interactive drama. In *Proceedings of the technologies for interactive digital storytelling and entertainment (TIDSE) conference*, volume 3, pp. 1–11.

N. Szilas. 2008. The mutiny: An interactive drama on idtension. In *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*, pp. 539–540.

P. Tambwekar, M. Dhuliawala, L. J. Martin, A. Mehta, B. Harrison, and M. O. Riedl. 2019. Controllable neural story plot generation via reward shaping. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pp. 5982–5988. AAAI Press.

B. Tearse, M. Mateas, and N. Wardrip-Fruin. 2010. Minstrel remixed: a rational reconstruction. In *Proceedings of the Intelligent Narrative Technologies III Workshop*, pp. 1–7.

M. Theune, S. Faas, A. Nijholt, and D. Heylen. 2003. The virtual storyteller: Story creation by intelligent agents. In *Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference*, pp. 116–127. Springer.

D. Traum. 2022. *Socially Interactive Agent Dialogue*, pp. 45–76. The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application. ACM, 2. DOI: http://dx.doi.org/10.1145/3563659.3563663.

S. R. Turner. 1993. Minstrel: a computer model of creativity and storytelling.

A. Tychsen, M. Hitchens, T. Brolund, and M. Kavakli. 2006. Live action role-playing games: Control, communication, storytelling, and mmorpg similarities. *Games and Culture*, 1(3): 252–275.

A. Tychsen, M. Hitchens, R. Aylett, and S. Louchart. 2009. Modeling game master-based story facilitation in multi-player role-playing games. In *Proceedings of the 2009 AAAI Symposium on Intelligent Narrative Technologies II*, pp. 24–32.

J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric. 2006. The role of physical embodiment in human-robot interaction. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 117–122. IEEE.

A. Weallans, S. Louchart, and R. Aylett. 2012. Distributed drama management: beyond double appraisal in emergent narrative. In *International Conference on Interactive Digital Storytelling*, pp. 132–143. Springer.

P. W. Weyhrauch. 1997. *Guiding interactive drama*. Carnegie Mellon University.

G. Zeglin, A. Walsman, L. Herlant, Z. Zheng, Y. Guo, M. C. Koval, K. Lenzo, H. J. Tay, P. Velagapudi, K. Correll, et al. 2014. Herb's sure thing: A rapid drama system for rehearsing and performing live robot theater. In *2014 IEEE International Workshop on Advanced Robotics and its Social Impacts*, pp. 129–136. IEEE.