This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1147/JRD.2019.2896157, IBM Journal of Research and Development

The long path to narrative generation

P. Gervás, E. Concepción, C. León, G. Méndez, P. Delatorre

Narrative generation, understood as the task of constructing computational models of the way in which humans build stories, has been shown to involve a number of separate processes, related to different purposes to which it can be applied, and focusing on specific features that make stories valuable. The present paper reviews a set of story generation systems developed at the NIL research group, each focusing on different aspects and functions of stories. These systems provide an initial breakdown of how the term storytelling might be either instantiated or broken down into component processes. The systems cover functionalities such as: generating valid plot structures, simulating character's behaviours or the evolution of affinities between them, either reporting or fictionalising events observed in real life, and revising a story draft to maximise the suspense it induces in its readers. These functionalities are not intended to exhaust the set of possible operations involved in storytelling, but they constitute an initial set to understand the complexity of the task. The paper also includes two proposals -- one theoretical and one technological -- for understanding how a set of such functionalities might be composed into a broader operational process that produces more elaborate stories.

Introduction

Storytelling is currently perceived as a fundamental component of the human cognitive ability [1] and as a crucial tool for successful participation in modern society [2]. It was also one of the earliest goals that Artificial Intelligence set itself when it first started trying to model human abilities [3] and remains an active field of research to this day [4]. In terms of the pragmatic definition of computational creativity as the "*engineering of computational systems which, by taking on particular responsibilities, exhibit behaviours that unbiased observers would deem to be creative*" [5], storytelling has a significant potential to contribute to the field.

From a pragmatic point of view, the task of modelling the human storytelling process constitutes an engineering challenge of the first order. Whereas successful prototypes of the storytelling ability have been achieved by employing individual technologies -- such as, for instance, planning [6], case-based reasoning [7] or agent-based modelling [8] to cite a few examples - it is unlikely that such a complex human ability can ever be modelled fully by recourse to a single technology. The fact that particular technologies lead to valuable results -- by capturing some of the features that make up a valid story -- clearly suggests that those technologies have something to contribute to the task. Yet the solutions in each case also show shortcomings with respect to features not specifically addressed by the technology in question.

The hypothesis underlying the material presented in this paper, is that each of these approaches focuses the modelling of storytelling on a particular subtask that is quite capable of producing stories with valuable features. Some of the features considered include the different views of stories as: narrative structures, simulations, evolving networks of characters affinity, narrations of observed facts, or suspense-driven entertainment. An important corollary of this hypothesis is that human storytellers rely on a combination of the subtasks based on these different features to obtain rich stories that combine the various features in rounded whole.

The present paper outlines a number of significant insights into the processes required to inject "narrative" quality into machine-enabled communication. The present section introduces the problem and describes the purpose of the manuscript. The second section outlines the different approaches that have been followed in storytelling research. The third section presents a number of alternative approaches for the consideration of stories, each one of them addressed in specific research systems to model different aspects of narrative communication. The fourth section addresses the challenge of integrating these subtasks into a single operational framework. This is done from two different points of view, one theoretical -- a conceptual analysis of narrative communication into a set of processes that interact in different ways to give rise to the different modes of communication described above-- and one computational -- the Afanasiev software engineering framework for allowing the composition of the computational systems described above.

Existing Approaches to Narrative Generation

There have been a significant number of efforts in the past to develop systems capable of generating stories, and an exhaustive review of them is beyond the scope of the present paper. Interested readers are invited to consult some of the existing surveys [3,4]. To provide a context for the set of different approaches considered here, this section reviews how the existing efforts in story generation have been categorised by prior authors.

An initial classification was provided by Paul Bailey [9], who considered four different approaches to storytelling, depending on the perspective under which stories were addressed: author models -- focusing on processes applied by authors --, story

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1147/JRD.2019.2896157, IBM Journal of Research and Development

models -- focusing on structural properties of stories --, world models -- focusing on world simulation -- and reader models -- focusing on the effect of stories on the reader. Gervás et al [10] proposed a classification of story generators based on the particular Artificial Intelligence technologies they employ, and identify generators based on planning and generators based on grammars. O'Neil [11] distinguishes between generators based on searching over the set of possible sequences of actions, and generators based on adapting existing stories into new versions to match a new purpose. Niehaus [12] establishes a distinction between systems that simulate a world from which the story is chosen, and systems that operate by deciding among a set of narrative elements such as plot structure, character dynamics, or the experience of the reader.

Kybartas & Bidarra [4] review systems from a wider perspective, allowing for collaboration between a human author and a computer system. In this context, they classify systems in terms of the degree to which the features under consideration are automated or left to be done manually by the human user. Within this view, they address as features *plot* -- the structure of the story -- and *place* -- the world that underlies the story.

In general terms, the variety of contrasting analysis suggests that there are many relevant aspects to story generation, and that different attempts at classification focus on specific ones while relegating others to secondary roles. However, successful modelling of human storytelling abilities is likely to require explicit treatment of all of these aspects in an integrated manner[13].

Deconstructing Storytelling

One of the ihe inspiring goals of research on storytelling has been to identify and develop mechanisms of interaction between people and machines that exploit language as means for human communication. Some of the solutions considered for this purpose have been based on Natural Language Processing, using various technologies for knowledge representation, and supported by a myriad of techniques from the domain of Artificial Intelligence. These mechanisms have been applied in practical contexts including intelligent information access, automated processing of medical documents, domotic interfaces, and verbal guidance for visually impaired users.

Although the empirical validation of these solutions was consistently positive, the solutions were often perceived as not reaching the level of fluency and naturalness that would be expected from a human carrying out the same tasks. However, it was difficult to identify what features were missing in the automated solutions that would have made them closer to human performance.

The progressive increase of public awareness of the importance of storytelling as a fundamental communication tool [1,2] suggested that the solutions that had been attempted were lacking a "narrative" quality that humans tend to impart to their communication efforts when they are directed to other humans. As a consequence, the research focus in the world of computational storytelling is progressively shifting to consider the role of narrative in communication, and to explore how this role might be modelled in computational terms.

However, the "narrative" quality that endears human-originated linguistic communication to human recipients is at the same time much less and much more than generating a story. Less because, in many cases, there is no need to invent a story to tell, but rather to "dress up" a given set of facts as a story, to use a known story as scaffolding for a content to be communicated, or to phrase a set of dry fact as if they were part of a story. More because, in other cases, it will involve inferring the best possible story implicit in a set of facts, or reading a story and finding a way to rephrase it that makes it more entertaining.

In this way, the challenge faced by a researcher hoping to build computational models of this kind of communication go beyond the generation of stories, and require understanding of how the stories are received, interpreted and validated. If stories are to be used as a code in a communication context, algorithms for encoding facts into stories need to be backed up with some notion of how the stories are decoded back into facts. For all of these processes, it becomes important to know what features are necessary for a discourse to be considered a story, and what makes a particular story "better" than another.

Addressing specific subtasks of storytelling

The work of modelling the human storytelling task has been addressed in the NIL research group less in terms of which existing technology might fare well as a single actor in the task and more in terms of which particular feature of stories should be given priority in a given implementation. As a result of this policy, several different systems have been developed.

Stories as Narrative Structures

One of the most visible component of stories is its narrative structure or plot. It is also a component on which there is a large corpus of theoretical work, both from the field of narratology and from the field of creative writing. Our most successful attempt to model narrative structure as a driving force for a story generator is the PropperWryter system.

PropperWryter is a program that generates the narrative structure for a single plot line, described in terms of a vocabulary of abstract representations of events that may happen in such a plot. It evolved from the Propper system [14], which generated plot structures for Russian folk tales based on Propp's Morphology of the Folktale [15]. Propp identified a set of regularities across a corpus of Russian folk tales in terms of character functions, understood as acts of the character, defined from the point of view of their significance for the course of the action.

The PropperWryter system is composed by a set of modules: a plot driver generator builds the sequence of Proppian character functions that determine the structure of the tale, a fabula generator instantiates the character functions in the resulting sequence with particular story actions, a casting module assigns particular characters to the arguments of the selected story actions, and a textual rendering module converts the final conceptual plan for a story into text.

An extension of the PropperWryter software as a generator of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1147/JRD.2019.2896157, IBM Journal of Research and Development

plot lines for musical theatre pieces was used for the composition of “Beyond the Fence”, the first ever experimental computer-generated musical. “Beyond the Fence” opened successfully at the Arts Theatre in London on 22nd February 2016 and enjoyed a successful two-week run [16].

Stories as Simulations

A different point of view on stories is to consider them as tools for communicating the evolution of a given world. To model this view, the STellA system (Story TELLing Algorithm) [17] builds stories from a set of simulations of what happens in the world.

STellA was designed as a take on the common sense knowledge bottleneck that constrains the possibilities of the existing story generation systems. Instead of focusing on one specific aspect of narratology or some concrete technology at hand to be tested for story generation, the design efforts in STellA were put on constructing a strong framework and an implementation for providing a knowledge-intensive solution.

In STellA, plots are assumed to be the result of very developed and complex interactions between characters and the world. As such, a detailed simulation takes place behind the scenes in order to produce a rich, complex material onto which to draw the narrative plot.

Therefore, two generative layers were carried out in STellA: first, the simulation layer applies physical, social and mental rules in a non-deterministic way. This means that many possible simulations are produced. The vast space of plots is a basic source of creativity, but this expansion is unconstrained and many non valuable stories are produced. The amount of “raw creativity” was approximated with a specific metric, the *entropy*, which reflects the amount of invented information in the simulation.

This exhaustive generation required several models running sequentially (and non-deterministically, for those models allowing for it): a physical ground for objects, collisions and movement, an agent layer for planning, a layer for describing emotions, grabbing things, attacking, loving, etc. While most of these layers were implemented according to not very complex models, connecting them into a logic process required a relatively high effort.

All this detailed simulation was enough to provide many coherente action sequences between characters and the environment surrounding them, but these sequences of events do not make an interesting narrative by themselves. STellA needed a new level of control over the generation in order to constrain the exploration and let the user set a number of input parameters for the system to produce a particular kind of narratives.

The identification of valuable plots was carried out by the *narrative layer*, which received partial simulations, discarded the non-promising ones, and set parameters for subsequent expansion. The narrative layer used explicit constraints, user objectives and *narrative curves* [18] to drive story generation. These curves measure and restrict the amount of suspense, action or entropy a simulation has introduced. **Figure 1** shows an example of curve evolution in STellA.

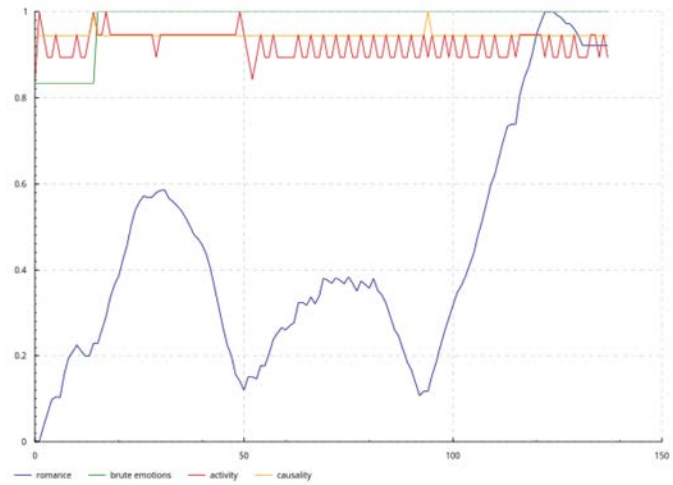

Figure 1 Curve evolution in STellA. As the story advances, the curves represent its narrative features.

The design of the knowledge base did not only affect the exploration of states, but the development of the narrative layer itself. Providing models for specific curves, especially in terms of the knowledge that STellA used to describe the plots, is a hard engineering problem in its own right. Identifying the amount of suspense in a certain point of the plot, the perception of danger or the evolving romanticism between a couple required both the identification of plausible models and the implementation in terms of the underlying representation.

STellA was successful and the implementation was capable of generating a massive amount of semantically valid stories. The costly engineering process of putting together many subsystems, plus the design challenge of a big architecture for non-deterministic generation is high, but STellA was able to produce a huge set of plots.

On the other hand, the exhaustive exploration was an intractable problem. The amount of plots that the system could generate was big, but the real problem is that there are no models for identifying when an *optimal* story was produced. Validity, coherence, user objectives and a shallow approach to narrative quality were implemented, but a real metric for *value* or *novelty* was not achieved.

Stories as Evolving Networks of Character Affinity

Stories may also be considered as reports on the evolution of a set of characters and the relations between them. The Charade system addresses this way of understanding stories.

Charade is a storytelling system that relies on a character simulation in order to generate the stories [19]. Charade builds upon Goethe’s theory of *elective affinities* [20] to represent human relations, which shows how affinities between the characters of a story can be represented by a topological chart. This theory was subsequently explored by Polti [21], where he works out the assertion made by Gozzi (the author of *Turandot*) saying that there could only be thirty six dramatic situations. After analyzing these situations, their variations and the characters involved, Polti concludes that the subtle differences

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1147/JRD.2019.2896157, IBM Journal of Research and Development

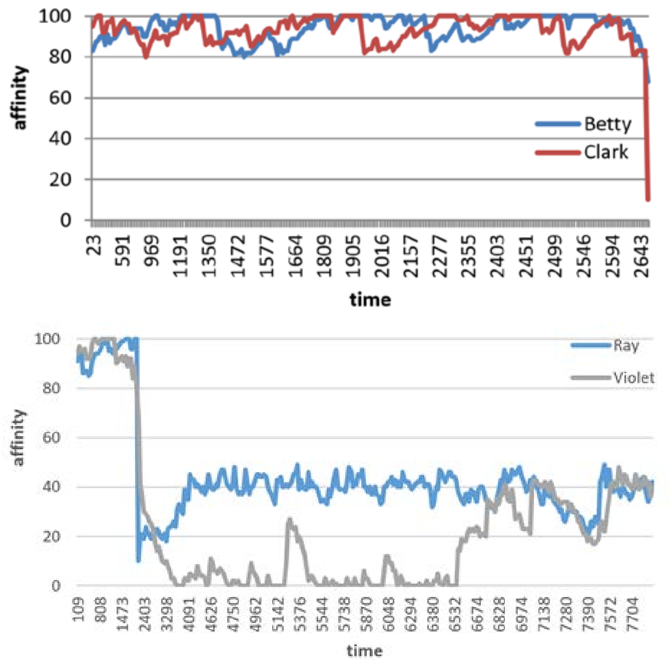

Figure 2 Evolution of the affinities between different characters in the Charade system.

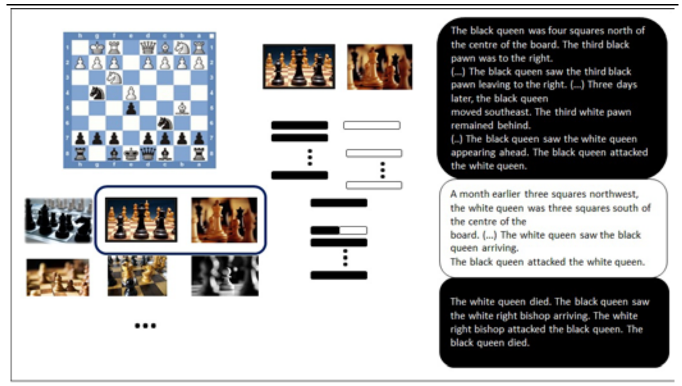

Figure 3 Operation of the Raconteur system. The description of a chess game from a file in algebraic notation is parsed into a set of narrative threads for individual pieces -- each covering a limited range of perception around the piece . A subset of these narrative threads is selected as focalised views of the story -- the two queens in this example. The chosen threads are spliced into a single discourse, adding statements to contextualise shifts in time and/or space as needed.

between these situations have to do with "*the ties of friendship or kinship between the characters*". A similar approach is used in *Thespian* [22], the social behaviour framework used to develop the *TLCTS* system [23].

The main objective of the Charade system is to study how character affinities can be used in storytelling systems in order to endow characters with believable interactions that can be used to enrich certain episodes within the plot. In order to achieve this goal, we have modelled four different levels of affinity: foe (no or low affinity), indifferent (slight affinity), friend (medium affinity) and mate (high affinity). We have used a fuzzy approach in order to model affinities, similarly to what is done in other cognitive or emotional architectures [24,25]. The advantage over a crisp model is that it allows us to overlap the affinity levels on their limits, which prevents relationships from changing constantly when moving around the limits of two different levels.

An additional feature of this affinity model is that it is not symmetrical. For any two characters, their mutual affinity is likely to have different values and it may even be situated in different levels, with the exception of mates. However, if they are not mates, one character may think that another character is a friend, while this second character may think the first one is a foe.

There are two mechanisms that make the affinity values change. The first one is the lack of interaction between two characters; in this case, the affinity values move towards the indifferent level. The second one is through interactions among characters, which obey a few simple rules. Each affinity level encompasses a set of possible interactions, so when dealing with another character, only actions pertaining to the affinity level between the two can be used. Otherwise, characters ignore interactions that are not contained in their perceived affinity level, and receiving such proposals penalizes the affinity with the character proposing them. Foes constitute an exception to this rule, as they carry out whatever they intend to do irrespective of the other character's intentions.

When receiving a proposal to do something together, a character may decide to either accept or reject it. If the proposal is accepted, both characters increase their mutual affinity values. If it is rejected, the proposer penalizes its affinity with the receiver. Actions for the same level of affinity have a different impact on it. For example, a romantic dinner has a stronger influence on affinity than going to the cinema together. Similarly, the negative effect of rejecting an invitation is opposite to the positive effect of accepting it. Two sample evolutions of the affinities between characters can be seen in **Figure 2**.

Stories as Narrations of Observed Facts

It is important to keep in mind that stories, regardless of the specific features that we have come to associate with them as value-added factors, originate essentially as tools for communicating a given state of the world, or the changes that affect it. The systems described in this section address this specific role of stories as vehicles for communicating a particular take on observed reality.

When the task of a storytelling author is considered there are several possible ways of viewing it depending on what the sources and the purpose of the intended story is. Although storytelling systems in the past had considered mostly the task of generating fictional stories, the systems described in this section consider the generation of stories either to convey facts observed from the real world, or to generate stories inspired by facts of this type.

The RACONTEUR system composes discourses to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1147/JRD.2019.2896157, IBM Journal of Research and Development

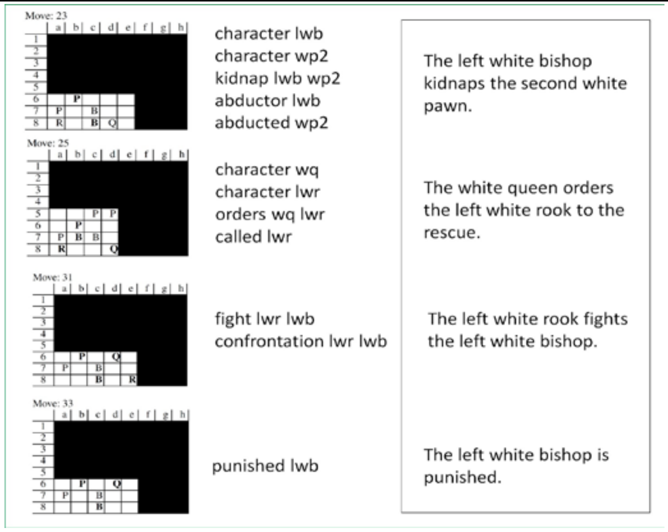

Figure 4 Example of a story produced by the StoryFire system. The narrative threads obtained for a given chess game by the Raconteur system are matched with a set of plots -- mined from Booker [34] -- based on presence of relevant characters as given for a particular mapping between chess pieces present at each point in the thread and narrative roles to be instantiated at each point in the plot. For a given alignment between thread and plot, a rendering as a story is produced relying on Raconteur functionality.

communicate (a selection of) the set of facts in a chess game [26]. The composition process operates as a self-evaluating cycle, in which the discourse that has been constructed at each point is decoded into a description of the facts it should communicate, and its quality measured in terms of how the interpreted version of the facts compares with the original. The operation of the system, which involves phrasing the events to be considered from the point of view of a participating agent -- a narrative thread centred on a particular character --, selecting a set of these, and then splicing them into a single discourse, is described in **Figure 3.**

This self-evaluating cycle can be understood as a baseline implementation of the reviewing stage of the writing task, as described by Flower & Hayes [27]. It can also be considered as an implemented instance of the ICTIVS model described later in the paper.

The StoryFire system [28,29] models the process of "fictionalising" real life events, whereby a set of facts that inspire a story, but either lack the structure or the clear motivation for the characters that one would expect in a story, are enriched or adapted until they coalesce into a surface form that exhibits the properties that make humans consider them a valid story. The computational model of the task of storifying a set of observed events involves: identification of a focalised narrative thread involving the events -- as in the Raconteur system --, selecting a plot structure that partially matches the chosen thread, mapping characters in the thread to characters in the plot, and generating a readable version of the resulting discourse. An example of a story generated in this way from a selected subset of moves of a chess game is given in **Figure 4.**

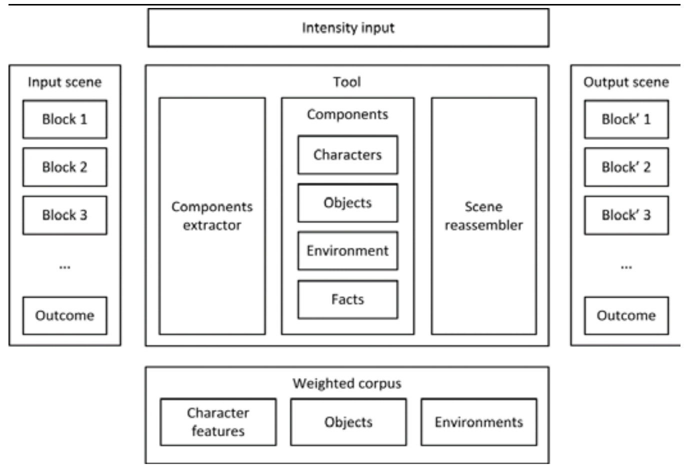

Figure 5 Architecture for rewriting scenes to maximise suspense.

Stories as Suspense-Driven Entertainment

A key feature of stories in our society is that they are produced not only for communication but also for entertainment. The entertainment value of stories is often determined by the suspense that they induce on the reader. This section describes a system developed to model this characteristic.

Delatorre et al. [30] propose the architecture of a system whose main objective is the adaptation of the descriptive elements of a scene, in such a way that the amount of information of the scene output is adjusted to the required suspense intensity. This architecture is represented in **Figure 5.**

The system receives as input a given scene -- represented in terms of a sequence of blocks followed by an outcome -- and a number indicating the desired intensity of suspense. From the input scene it extracts a number of components -- characters, objects, environment and facts -- which it then analyses based on a weighted corpus in which each concept is associated to a quantitative value that represents its level of suspense. Examples of relevant features to establish this value have been shown to be: for characters, empathy and relative strength of individuals in conflict, or for objects, their functional nature that might affect the outcome -- as in weapons or possible exits. The values were empirically obtained from experiments where human participants were asked to rate variations of a given story controlled for contrasting instantiations of the features [31]. Based on a computed overall value for suspense at each point arising from the various concepts present, the system then reassembles the scene -- replacing concepts with differently valued alternatives -- in order to achieve the desired level of suspense.

Relation to prior classifications

Each of the systems described so far presents characteristics that would classify it differently along the taxonomies established by prior researchers, which were reviewed above in the section **Existing Approaches to Narrative Generation**.

The PropperWryter system might be classified along the lines of Bailey's story models, Niehaus's systems that decide based on narrative elements, or Kybartas & Bidarra systems focusing

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1147/JRD.2019.2896157, IBM Journal of Research and Development

on plot.

In contrast, the essential focus on simulation followed by the STellA system would place it more along Bailey's world models, Niehaus's world simulators, or Kybartas & Bidarra systems focusing on place. Yet in this case, the use of narrative curves as a decision mechanism includes elements that would fall under Bailey's story models, or Kybartas & Bidarra systems focusing on plot.

The Charade system relies on world models at a different level of representation, more focused on characters, and in that sense would fall under Bailey's world models and Kybartas & Bidarra systems focusing on place.

The Raconteur and StoryFire systems have a fundamental ingredient of modelling the author's task which would classify them as Bailey's author models, focused on a very specific version of the storytelling task that addresses the distinction between factual and fictional stories. But the fact that they start from input in terms of a world model would mean that they also be considered as a reader model. In a sense, this system establishes a bridge between world models and story models, in as much as it encodes a procedure for connecting a view of the world with a representation of it as a story, and one that takes a model of the reader into account to inform the process.

The suspense-driven scene rewriter system is also addressing Bailey's author models, in the sense that it focuses on generating suspense for entertainment purposes. However, it also classifies as a reader model in the sense that its whole goal is to produce a particular effect on the reader.

Integrating subtasks into an account of storytelling

It is clear that the different systems described so far address features of stories that are relevant to their perceived quality, and that they emulate different functionalities that human storytellers bring to the task. Each of the systems constitutes a first approximation at modelling how that particular functionality contributes to storytelling. This poses two interesting questions. First, further understanding of how these functionalities cooperate together to the generation of stories is required. Second, some means for combining together the existing computational solutions for these subtasks would reduce the cost of developing solutions.

Key processes in narrative communication

The field of storytelling would benefit greatly from a theoretical model of the task as a composition of simpler processes. Such a model might propose protocols or procedures for how the various functionalities outlined above contribute to one another and to the overall shape of a story. Researchers at the NIL group have attempted to put together a pragmatic description of how some of the tasks that had been identified as part of the group's research into an operational model that articulates them into a computational process. This model is not intended as a description of how humans address the storytelling task, but rather as a set of interactions between subtasks that might replicate some of the behaviour observed in humans.

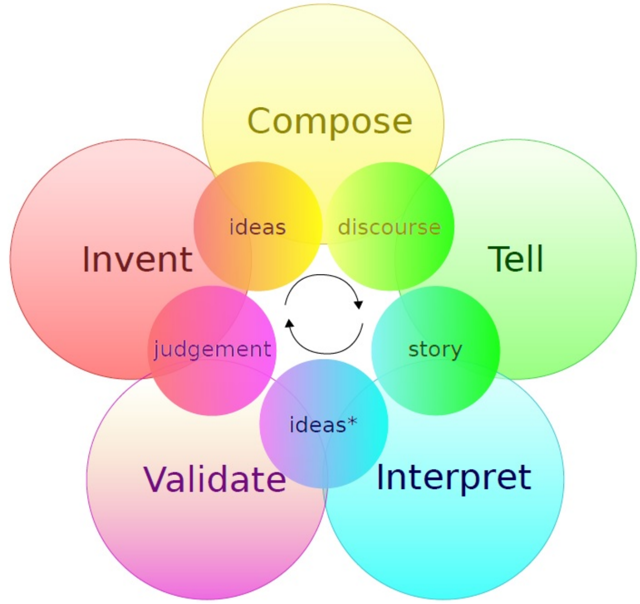

Figure 6 Illustration showing the potential cycle through the stages of the ICTIVS model. In particular cases, the task may take place involving only some of the stages.

The different kinds of communication between people that can be considered to have a "narrative" quality involve a complex feedback cycle that may include tasks for inventing content, for organising content, for interpreting content, and for validating content. Instances of this type of cycle may include all possible tasks (when an author invents a plot, revises it, develops the discourse to tell it, then revises that) or only some of them (when someone processes the memory of the day at work into a story to tell the family on arriving home -- which is unlikely to involve any invention of content, but may include careful consideration of what they are going to infer from what is told --, or when a child invents a story as he plays -- which is unlikely to include revision of any kind). A theoretical description of this set of tasks has been postulated in the ICTIVS (Interpretation, Composition, Telling, Interpretation and Validation of Stories) model [13].

In order to capture separately the tasks of coming up with the content to be transmitted, encoding it into the discourse that will form the message, and consider the process of tentative interpretation implied in its validation, the model includes five stages: *INVENTION* -- creating or establishing the content for the message; *COMPOSITION* -- constructing the discourse that conveys the message; *INTERPRETATION* -- applying mirror processes to those that will be applied by the receiver to estimate what will be understood; *VALIDATION* -- predicting the impact that the message as sent, via the interpreted constructed from it, may have on the receiver; and *TRANSMISSION* -- actual operation of sending the message to the receiver. The sender may cycle over the first four stages until satisfied that the interpretation and the impact predicted by her own mirror processes correspond to the reaction she expects from the receiver, and only then will she decide to transmit the message. This potential cycle though the stages is represented in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1147/JRD.2019.2896157, IBM Journal of Research and Development

Figure 6.

Microservice Architecture for Combining Computational Storytelling Subtasks

The existence within the group of several systems implementing different computational solutions for the storytelling tasks, and the insights described above concerning the need to find ways of integrating the different tasks into a single operational process, lead to attempts to provide efficient computational solutions for combining together different storytelling functionalities.

Afanasyev [32] is a microservice-based architectural framework for enabling the construction of distributed story generation systems. The core concept behind its model is allowing the existing story generation systems to break their functions into independent services and integrate them in a collaborative environment for generating stories. The main advantage of this approach is the composition of diverse generation models through easily replaceable microservices.

Promoting the interoperability between such diverse components entails the development of an associated shared representation model [33]. This model not only considers a mere syntax formalism, but also a knowledge representation model as a base for a collaborative environment to run an enhanced process of literary creation. This representation model requires a knowledge base that can be fed with data from the diverse storytelling systems.

Also, the addition of an pre-existing service usually entails the development of an Afansyev's signature-compliant wrapper, due to the need of adapting it to the framework-specific representation model.

Every service of the reference architecture fits a predefined role, as depicted in **Figure 7**. Every role is formalized by a REST-based signature that must be implement by the microservice that carries it into execution.

The component that drives the whole ecosystem is the Story Director. Its primary goal is to orchestrate the activity of the different microservices when performing the story generation process. To this end, it invokes systematically the REST APIs defined in the framework and provided by the microservices. The behavior of the Story Director is provided according to the Strategy pattern approach, so it can be easily replaced by a different model, and thus the generation process.

The Plot Generator in Afanasyev is responsible for creating the basic structure of a story, namely the plot. It creates a skeleton containing the relevant episodes that happen during the story, and the preconditions and postconditions related to every episode. Afanasyev's model allows for maintaining several plot generators that will be conveniently selected by the Story Director to create the plot of a story.

The Episode Generator is the microservice that works in the completion of every episode defined in the plot. It tries to fulfill the preconditions and postconditions established for the episode.

The Filter Manager applies a sequence of predefined filters over every episode for guaranteeing the quality of the generated story in terms of drama, tension, suspense or any other

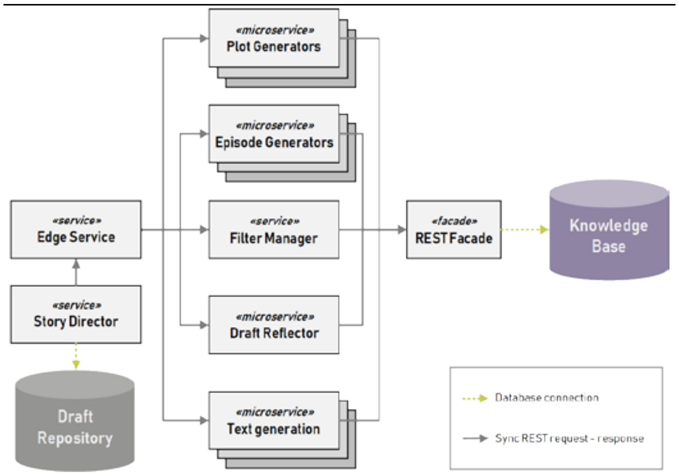

Figure 7 Graphical depiction of the Afanasyev's reference architecture

interesting parameter. If the episode is rejected, then the Story Director requests the Episode Generator to generate a new one.

The Draft Reflector reviews globally the ongoing draft, that is, the story that is being created, and determines if it has been concluded and must be considered a story, or if requires more iterations for completing the episodes that have not yet been developed in the draft.

Conclusions

The attempts to develop computational models of storytelling have shown that, far from being a well-defined task with clearly specified inputs and outputs, what people understand by the concept of storytelling is in fact a set of considerably diverse operations that are sometimes carried out in isolation to achieve simple stories or specific ingredients that might be a part of stories, and sometimes combined into the production of more elaborate stories.

The systems presented in this paper constitute a small set of examples of computational implementations of instances of these operations. They are not exhaustive in their coverage of processes that might be involved in storytelling. Indeed, there surely exist many others that are equally fundamental -- or even more so -- and which have yet to be modelled.

The main insight arising from the analysis of this collection of systems is that research efforts in computational storytelling should stop describing themselves under broad and vague labels, and should start being more specific as to which particular feature of stories they focus on, what purpose they build their stories for, and what sources of material they are concerned with including in their stories.

The present paper outlines a number of possible descriptions that constitute a first approximation of a vocabulary of constituent tasks of the general storytelling ability. This vocabulary is very much in need of further enrichment if it is to cover the fascinating complexity of the task as carried out by humans.

Another important insight is that models are required of how these constituent tasks of storytelling interact with one another to give rise to the rich stories and complex storytelling processes that we see in human authors. The paper describes two working models -- one theoretical and one technological -- as initial approximations of how to solve this problem. Nevertheless, this is clearly an open problem that will need significant further work.

Acknowledgments

This work has been supported by the IDiLyCo project (TIN2015-66655-R) funded by the Spanish Ministry of Economy, Industry and Competitiveness; and by the project FEI INVITAR-IA of the Universidad Complutense de Madrid.

References

1. Bruner, J. The Narrative Construction of Reality, Critical Inquiry, Vol. 18, No. 1, pp. 1-21, (1991).

2. Salmon, C., Storytelling: La machine à fabriquer des histoires et à formater les esprits, Découverte, (2007).

3. Gervás, P. Computational Approaches to Storytelling and Creativity, AI Magazine, Vol 30, No 3, (2009)

4. Kybartas, B. , Bidarra, R., A Survey on Story Generation Techniques for Authoring Computational Narratives, IEEE Transactions on Computational Intelligence and AI in Games, 9(3), 239-253 (2017)

5. Colton, S., Wiggins, G., Computational creativity: the final frontier? in: Proceedings of the 20th European Conference on Artificial Intelligence, Pages 21-26 (2012).

6. Riedl, M.O., Young, R.M. Narrative Planning: Balancing Plot and Character. J. of Artificial Intelligence Research 39, 217–268 (2010).

7. Turner, S.. MINSTREL: A Computer Model of Creativity and Storytelling. PhD thesis, University of California at Los Angeles, Los Angeles, CA, USA, (1992).

8. Theune, M. Faas, E., Nijholt, A., Heylen, D. The virtual storyteller: Story creation by intelligent agents. In Proc. of the Technologies for Interactive Digital Storytelling and Entertainment, pp 204–215, (2003).

9. Bailey, P. Searching for Storiness: Story-Generation from a Reader's Perspective. In Narrative Intelligence. AAAI Tech. Report, FS-99-01, AAAI Press, Menlo Park, California, 157-163.(1999).

10. Gervás, P., Lönneker-Rodman, B.,Meister, J.C., Peinado, F., Narrative models: Narratology meets artificial intelligence. In R. Basili and A. Lenci, editors, International Conference on Language Resources and Evaluation. Satellite Workshop: Toward Computational Models of Literary Analysis, pages 44–51, Genova, Italy, 2006.

11. B. O'Neil. A Computational Model of Suspense for the Augmentation of Intelligent Story Generation. PhD thesis, Georgia Institute of Technology, Atlanta, Georgia, USA., (2013)

12. J. Niehaus. Cognitive Models of Discourse Comprehension for Narrative Generation. PhD thesis, North Carolina State Univ, (2011)

13. Gervás, P., León, C., "The Need for Multi-Aspectual Representation of Narratives in Modelling their Creative Process", in 2014 Workshop on Computational Models of Narrative, Canada,(2014)

14. P. Gervás, "Computational Drafting of Plot Structures for Russian Folk Tales", Cognitive Computation (2015).

15. V. Propp: Morphology of the Folktale. 1968.

16. S. Colton, Llano, M. T., Hepworth, R., et al., "The Beyond the Fence Musical and Computer Says Show Documentary", in Seventh International Conference on Computational Creativity (ICCC 2016), Paris, France, (2016).

17. León, C., Gervás, P.: Creativity in Story Generation From the Ground Up: Non-deterministic Simulation driven by Narrative. In: 5th International Conference on Computational Creativity, ICCC 2014, Ljubljana, Slovenia. (2014)

18. León, C., Gervás, P.: Prototyping the Use of Plot Curves to Guide Story Generation. In: Workshop on Computational Models of Narrative, Language Resources and Evaluation Conference (2012)

19. Méndez, G., Gervás, P., and León, C.: On the Use of Character Affinities for Story Plot Generation. In: Knowledge, Information and Creativity Support Systems: Selected Papers from KICSS'2014, vol. 416, pp. 211-225, Springer. (2016)

20. von Goethe, J. W.: Elective Affinities / Kindred by Choice. (1809)

21. Polti, G.: The Thirty-Six Dramatic Situations. (1917).

22. Si, M., Marsella, S.C., Pynadath, D.V.: Thespian: Modeling socially normative behavior in a decision-theoretic framework. In: Intelligent Virtual Agents, Lecture Notes in Computer Science, vol. 4133, pp. 369-382. Springer (2006)

23. Johnson, W.L., Valente, A.: Tactical language and culture training systems: Using artificial intelligence to teach foreign languages and cultures. In: AAAI. pp. 1632-1639 (2008)

24. El-Nasr, M. S. , Yen, J., Ioerger, T.R.: Flame - fuzzy logic adaptive model of emotions. In: Autonomous Agents and Multi-agent systems, 3(3):219-257. (2000)

25. Imbert, R., de Antonio, A.: An emotional architecture for virtual characters. In: International Conference on Virtual Storytelling, pages 63-72. (2005)

26. Gervás, P. "Composing Narrative Discourse for Stories of Many Characters: a Case Study over a Chess Game", Literary and Linguistic Computing, vol. 29, (2014)

27. Flower, L., Hayes, J.R. A cognitive process theory of writing. College Composition and Communication, 32(4):365–387, (1981).

28. Gervás, P. "Storifying Observed Events: Could I Dress This Up as a Story?" in 5th AISB Symposium on Computational Creativity, University of Liverpool, UK, (2018)

29. Gervás, P. "Targeted Storyfying: Creating Stories About Particular Events", in Ninth International Conference on Computational Creativity, ICCC 2018, Salamanca, Spain, (2018)

30. Delatorre, P., Arfè, B., Gervás, P., Palomo-Duarte, M.: A component-based architecture for suspense modelling. In: Proceedings of AISB 2016's Third International Symposium on Computational Creativity. (2016)

31. Delatorre, P., León, C., Gervás, P., y Palomo-Duarte, M.: A computational model of the cognitive impact of decorative elements on the perception of suspense. Connection Science, 29(4):295–331. (2017)

32. Concepcion, E., Gervas, P., Mendez, G.: Afanasyev: A collaborative architectural model for automatic story generation. In: 5th AISB Symposium on Computational Creativity. AISB 2018. (2018)

33. Concepcion, E., Gervas, P., Mendez, G.: A common model for representing stories in automatic storytelling. In: 6th International

Workshop on Computational Creativity, Concept Invention, and General Intelligence. C3GI 2017. (2017)

34. Booker, C. The Seven Basic Plots: Why We Tell Stories. Continuum, (2004).

Pablo Gervás *Facultad de Informática, Universidad Complutense de Madrid, Spain (pgervas@ucm.es)*. Dr. Gervás is Associate Professor at the Department of Software Engineering and Artificial Intelligence of the School of Informatics at the Universidad Complutense de Madrid. He is the director of the NIL research group (Natural Interaction based on Language). Over the years, his research interests have shifted towards studying the role of narrative in human communication, with a view to applying it in human-computer interaction. Besides his generic interest in understanding and modelling language, he tries to tackle computational models of human creativity. His interests include the following lines of research: Natural language generation, Natural language analysis, NLP for accessibility, and NLP and literary artifacts.

Eugenio Concepción *Universidad Complutense de Madrid, Madrid, Spain(econcepc@ucm.es).* Concepción received his M.S. degree at Universitat Oberta de Catalunya (UOC) and B.S. degree in Computer Engineering at Universidad Complutense de Madrid (UCM). He is currently developing his PhD in Computer Science at Universidad Complutense de Madrid (UCM) focused in Automatic Story Generation. He combines this research activity with his responsabilities as CTO in a software development company.

Carlos León. *Facultad de Informática, Universidad Complutense de Madrid, Spain (cleon@ucm.es)*. Dr. León received his PhD in Computer Science from the Universidad Complutense de Madrid (UCM) in 2010. He is currently an Assistant Professor in the Department of Software Engineering and Artificial Intelligence of the Complutense University of Madrid. Dr. León's research focuses on computational models of creative behavior, with particular focus on cognitive descriptions of narrative generation. He studies the role of narrative as a fundamental component of human cognition.

Gonzalo Méndez *Facultad de Informática, Universidad Complutense de Madrid, Spain (gmendez@fdi.ucm.es)*. Dr. Méndez received his PhD in Computer Science from the Universidad Politécnica de Madrid (UPM) in 2008. He is currently an Associate Professor in the Department of Software Engineering and Artificial Intelligence and Director of the Instituto de Tecnología del Conocimiento at the Universidad Complutense de Madrid. His current research interests include narrative generation, natural language generation for figurative language, accessibility systems and agent-based simulations.

Pablo Delatorre. *Facultad de Informática, Universidad de Cádiz, Spain (pablo.delatorre@uca.es)*. Prof. Delatorre received his M.S. degree and B.S. degree in Computer Engineering at Universidad de Sevilla. He is currently developing his PhD in Computer Science at Universidad Complutense de Madrid (UCM), focused in Narrative and Interactive Suspense applied to Automatic Story Generation. He combines this research activity with his responsibilities as teacher at University of Cadiz. Outside the academic profession, he worked ten years in the ICT sector, managing national and international corporate projects.