Computational Models for Understanding Narrative

Modelos Computacionais para Entender a Narrativa

NICK MONTFORT

Massachusetts Institute of Technology, Cambridge, MA, USA
Center for Digital Narrative, University of Bergen, Norway
nickm@nickm.com

RAFAEL PÉREZ Y PÉREZ

Universidad Autónoma Metropolitana–Cuajimalpa, México
Center for Digital Narrative, University of Bergen, Norway
rperez@cua.uam.mx

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.58 (2023) ISSN 2183-7198

Abstract

We describe how the computational modeling of narrative serves as a method of inquiry and helps to further humanistic understanding in this domain. Our focus is on our own systems, MEXICA and Curveship. Each of these two computational narrative systems is a working representation of aspects of the human processes of creative writing or narrating, and can be used to explore these processes and learn more about them. We describe some specific insights gained regarding the connection between characters' emotional relationships and conflicts, collaboration between writers, elements of narrative theory, expressions of surprise, and how referring expressions are important to literary style. We conclude by considering how models of *story* and *narrative* are not the same as large language models (LLMs) and we should not expect either type of system to do the work of the other.

Keywords

cognition | engagement-reflection | narratology | plot| style

Resumo

Descrevemos como a modelagem computacional da narrativa serve como um método de investigação e ajuda a aprofundar a compreensão humanística neste domínio. O nosso foco está nos nossos próprios sistemas, MEXICA e Curveship. Cada um desses dois sistemas narrativos computacionais é uma representação funcional de aspetos dos processos humanos de escrita criativa ou narrativa e pode ser usado para explorar esses processos e aprender mais sobre eles. Descrevemos algumas visões específicas obtidas sobre a conexão entre relacionamentos emocionais e conflitos dos personagens, colaboração entre escritores, elementos da teoria narrativa, expressões de surpresa e como as expressões de referência são importantes para o estilo literário. Concluímos considerando como os modelos de história e narrativa não são iguais aos large language models (LLMs) e não devemos esperar que nenhum tipo de sistema faça o trabalho do outro.

Palavras-chave

cognição | engajamento-reflexão | narratologia | enredo | estilo

1. Introduction

Recently, the first book providing a historical overview of computer systems that do creative writing was published (Sharples & Pérez y Pérez 2022). That history deals with how different types of creative systems, ranging from popular ones to ones for academic research, have been developed over the years, often but not always in conversation with each other. In recent years, there have also been several good articles published that survey storytelling systems, including (Gervás 2009), (Kybartas & Bidarra 2017), (Hou et al. 2019), (Herrera-González et al. 2020), (Alhussain & Azmi 2021). A book detailing the operation of story generators will be released this summer (Pérez y Pérez & Sharples Forthcoming).

We refer the interested reader to these resources for an overview of computer story generation and narrating. Our discussion focuses on just two systems, ones that we have developed. We also consider one main purpose that systems like this can have: How they can enhance our understanding by modeling different aspects of writing and narrative. Systems to generate plots can also do other things, including supporting creative work. But our focus in this discussion is their use, in research that relates to humanistic inquiry, as models. In many sorts of research, it can be fruitful to proceed by developing a theory about whatever we are trying to understand — the spread of disease in a population, predicting storms in a weather system, determining the effects of a different monetary policy in an economy, etc. We then would like to test our theory to see if it fits the world well. A computational implementation, in the form of a model, is one way of doing so, even in cases (such as the three listed) where we cannot easily conduct experiments in the real world.

In the discussion that follows, we describe Pérez y Pérez's MEXICA, which among other things is a model of the creative writing process (for developing the plots of stories, in particular), and Montfort's Curveship, a system for modeling narrative variation or narrative style. We also discuss two projects to integrate these systems and outline some specific insights produced by using MEXICA and Curveship. Then, we distinguish models of story and narrative from the natural language processing systems that are attracting the most attention currently, large language models (LLMs) and specifically the systems in the GPT family.

2. *The Point of Developing Models*

In the humanities we have helpful *theories* and *accounts* of different cultural phenomena, including writing and narrating. Building an operational, computational model of some humanistic theory (such as narrative theory) or some cognitive account allows us to inquire in several ways, for instance, helping us determine what these systematic models actually describe in detail and what they are missing. Those of us doing computational research not only *discuss* humanistic theories, we *operationalize* them and make them function as computer models. If we cannot implement aspects of these theories without further theorizing and elaboration, this suggests that the theories are underspecified in some ways. Of course, there could be other explanations for our inability to computationally model humanistic theories: The researchers attempting it may not be up to the task. Over time, however, we would expect that in the entire research ecology, those theories that are operationalizable will come to be modeled, whether by the researchers who originally attempted it or by others. After modeling is done, computational models based on these theories and accounts also allow us to investigate how well they fit with whatever phenomena they are trying to explain.

Our discussion here covers two main preexisting theories. The Engagement-Reflection cognitive account of creative writing (*E-R Cognitive Account*) offers a general description of how we write. *Narratology* or narrative theory is among other things a description of how narrating can be done. In both cases, developing computer systems to embody these theories requires a very detailed account of the processes of writing and narrating. In a metaphorical way, we claim that in the construction of a computational model of a cognitive or social phenomenon, there is a struggle between those forces that give meaning to a general description, and those forces that demand that details of the processes involved in the experience be represented algorithmically (Pérez y Pérez 2018). To resolve this tension, it is necessary to build a bridge that allows transit between the general ideas and the concrete implementation. That is the role of a computer model. This is also why we distinguish between the E-R Cognitive Account and the E-R Computational Model of MEXICA, and also between narratology (including the narratology of specific theorists) and the particular model of narrative variation implemented in Curveship. The former cognitive descriptions provides a framework to build on and advance the latter computational models.

3. *MEXICA*

Pérez y Pérez (1999) developed a computational model of the creative process that implements the pre-existing "Engagement-Reflection Cognitive Account of Creative Writing" (E-R Cognitive Account). The computational system is MEXICA (Pérez y Pérez and Sharples 2001, 2004; Pérez y Pérez 2007), and produces narratives about the Mexicas, the ancient inhabitants of the Valley of Mexico, often called the Aztecs (Pérez y Pérez 2017). The most recent version of MEXICA is also able to produce narratives in different settings.

The E-R Cognitive Account is based on ideas expressed by different researchers that were collected and extended by Mike Sharples and describe how the creative process works when we write (Sharples 1999). Sharples's concepts can be summarized as follows: The creative process consists of a constant cycle between two mental states known as Engagement and Reflection. During Engagement people are immersed in the generation of sequences of new ideas through associations: an idea produces a context that leads us to associate another new idea, which leads to another new one, and so on. A typical example is daydreaming, where ideas just flow and we have no control over them. Engagement is interrupted when we are distracted or when we get blocked and cannot generate more material. Then, we switch to Reflection, where we evaluate and, if necessary, modify the material generated so far. This reflective evaluation produces a series of guidelines that condition the generation of new material during Engagement. Once the evaluation is completed, we return to Engagement and the cycle continues.

Often in thinking about writing, we consider the difference between producing a draft and revising it. These two activities have some relationship to Engagement and Reflection, but they happen over a much longer span of time. The E-R Cognitive Account describes types of thinking that happen minute-by-minute, whether we are drafting or revising.

The E-R Cognitive account was the framework to build the E-R Computer Model. The program MEXICA is an instantiation of this model. The main goal of MEXICA is the generation of narratives that are novel, coherent, and interesting. In MEXICA, a story is defined as a sequence of actions, not a surface text. A story is novel if it is not similar to the stories that the system has stored in its knowledge base. A story is coherent if its actions fulfill common-sense knowledge. MEXICA keeps track of events in the story that generate tension; thus, a story is interesting when it includes variations in the story's tension.

Two characters in love, or hating each other, provide examples of emotional links. Conflicts modeled by MEXICA include a character being injured or deprived of freedom.

MEXICA uses a database which represents its knowledge. This database is made up of a set of records, called knowledge structures, which are organized in terms of emotional links and conflicts between characters. Each of these structures is associated with logical actions to be executed. For example, the database might record that, when two characters are in love (an emotional link), a coherent way to continue a narrative is that these characters get married, or go on a trip together, or move in together. Similarly, the database might record that, when a character is injured (a conflict), among the logical actions to continue the story are that the character finds a way to treat the injury, another character helps, or the injured character dies. For details on how the database is built, see (Pérez and Pérez 2007).

In MEXICA, all the actions that are carried out in a narrative have a set of consequences associated with them. These consequences are in terms of emotional links and conflicts between characters. For example, consider when the system generates an event in which the knight kidnaps the princess. The princess becomes a prisoner (a conflict) and she also begins to hate the knight (an emotional link). Since both characters are in the same place, a new conflict known as potential danger is triggered, which represents that the physical integrity of the knight is in danger, because the princess's hatred may lead her to attack him. This information is stored in a structure known as the story context. In this way, as the plot develops, the context of the story becomes more elaborate.

To generate a new narrative, the user provides an initial action. The system executes this action and the story context is generated. Next, the system searches the database for a knowledge-structure that is the same or similar to the story context and retrieves its set of possible actions to continue the story. MEXICA chooses one of these possibilities at random, executes it, the story context is updated, and the cycle repeats.

After generating a sequence of three actions (the number generated is a parameter which can be modified), the system switches to the reflection state. There, the system checks if the story in progress is interesting, coherent, and novel. If any of these attributes has a low evaluation, the program employs a series of heuristics designed to improve them during the next cycle of engagement. The system then returns to start a new Engagement cycle. If, during Engagement, the system does not find options to continue the plot, it switches to reflection where it inserts an action to try to break the impasse. The goal is that this new action produces a story context which can be matched with some structure in the database. If the impasse cannot be broken, MEXICA resolves those conflicts that are still active and ends the narrative. For example, if a character is injured, the program inserts requisite actions so this character either recovers or dies.

MEXICA has a series of parameters that allow modifying the pace with which a story unfolds. In this way, a narrative developed by MEXICA can be the result of a process that includes several cycles of engagement and reflection, or a process with a clear tendency towards the state of engagement and little reflection, or a process oriented towards the state of reflection and little engagement. Of course, the content of the database also affects this behavior.

MEXICA was conceived as a research tool; therefore, anyone interested in using it can define a significant number of parameters that control the operation of the system. Any computational model can be modified more substantially by undertaking additional software development, which has been done throughout the lifetime of the system.

This is an example plot produced by MEXICA, represented here using English templates to generate sentences:

> As part of a conspiracy, the eagle knight took the lady hostage, planning to kill her.
> The lady could not help it and she humiliated the eagle knight.
> Striking quickly, the eagle knight injured the lady.
> Angry, the lady attacked the eagle knight!
> Fiercely, the lady hurt the eagle knight.
> The lady made a potion and drank it quickly. She started to recuperate!
> The eagle knight went to find some medicinal plants and cured the eagle knight.
> He was lucky!
> The eagle knight went to Texcoco lake.
> Quietly, the lady was able to escape!
> The end.

This plot was developed as follows. The user of the system provided an initial action: The eagle knight kidnapped the lady. Then, the system started and, during Engagement, generated a sequence of three actions: The lady insulted the knight, the knight wounded the lady, and the lady reacted by attacking the knight. MEXICA did not find a way to continue the tale. So, it switched to reflection and inserted the action in which the lady injured the knight as a way to try to break the impasse. It switched back to engagement, but still, it could not find a good association connecting the new events. So, the system decided to end the tale and tried to sort out all the unresolved conflicts in the story. Thus, MEXICA inserted actions in which the lady and the knight cured themselves, in which the knight decided to run away (clearly, they could not be together without hurting each other), and in which the lady regained her freedom. Then, MEXICA evaluated the narrative and generated a report. The following is part of that report:

— The beginning is too abrupt.

— The story incorporates situations where a character interacts with itself in an unconvincing way. For instance: THE EAGLE KNIGHT WENT TO FIND SOME MEDICINAL PLANTS AND CURED THE EAGLE KNIGHT. HE WAS LUCKY!

— The narrative opens with an introduction where each event clearly connects to the next one and works as a foundation for the rest of the tale.

— The tale reaches its climax in the scene where "FIERCELY, THE LADY HURT THE EAGLE KNIGHT".

— This story is wonderful!

— My evaluation of your story is ->95/100

The report points out that, rather than developing a conflict step by step, the story starts with a kidnapping situation. The evaluation indicates that the tale includes an action where a character interacts with himself in an unconvincing way. As a result, the final score is decreased slightly. The report describes that actions have a clear cause-effect relation and a climax is reached when the lady injures the knight. The story satisfies the system's requirements and gains a final evaluation of 95/100.

4. *Curveship*

One of the important qualities of narrative is that regardless of what underlying events transpire in the story world, these events can be represented in many different ways — that is, the story level (or content) can be understood as distinct from the narrative discourse (or expression). A substantial body of theory about narrative, which considers narrating as a central issue and takes the story/discourse distinction as essential, has developed since the 1970s and is known as narratology or narrative theory.

Montfort developed a system, beginning in 2006, that came to be called Curveship. (Prior to 2011, the same system was called nn.) This system computationally models some of the most significant aspects of narratology (Montfort 2007, 2011). While it is convenient to speak of narratology as a field, there are many narratologies advanced by different theorists who agree about many, but not all, points. Curveship is mainly a model of the narratology of Gerard Genette (1983, 1988), also incorporating important ideas from Gerald Prince (1982) and Marie-Laure Ryan (1991, 2001).

In contrast to MEXICA, Curveship by itself does not model creativity, nor does it have a focus on cognition. Rather, it is meant to explore how a fairly small number of underlying parameters, and a reasonably simple model, can be used to accomplish many different types of narrating that have been observed and accounted for by narrative theorists. Curveship for example does have a model of actions that draws on the cognitive account of Conceptual Dependency theory (Schank 1972), but even this foundational aspect of the system is mainly incorporated to allow for effective narrating as described by narratology; the purpose of the system is not to test whether this account

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.58 (2023) ISSN 2183-7198

is cognitively valid. Because narratology theorizes how underlying events can be told in different ways, Curveship models these possibilities and variations. The system does not embody any notion that one way of narrating is better (for instance, more creative) than any other. Curveship does not take the position that some types of narrative are more authentic (representative of real human thought and activity) than others.

The literary work most related to Curveship is Raymond Queneau's *Exercises in Style* (1981), a collection of vignettes that all relate the same underlying and essentially uninteresting events, but do so in ninety-nine different ways. This lively work exhibits how, even if events at the story level are not that compelling, various ways of narrating those events can be extremely engaging. To make a fine distinction, Curveship's particular focus is not writing style in general, but *narrative style* — variations that are particular to the representation of events. So Curveship can reverse the order of events in the telling, beginning with the last one and proceeding through to the first one, as in Queneau's "Retrograde." But it does not attempt to model Queneau's styles that are based on slang, which could be used to relate greetings, describe things, and produce other sorts of texts just as easily as the style could be used to narrate.

While MEXICA is a plot generator, Curveship is essentially a text generation system, and follows the classic three-stage model of document planning (determining what content to include), microplanning (making specific lexical choices, including determining when referring expressions will be used), and realization (the final production of natural language), a model detailed in (Reiter & Dale 2000). It differs from other systems in that it exposes parameters, collectively called *spin,* which pertain specifically to narrative: To the representation of events by a narrator, more or less overt.

The original Curveship system is now called Curveship-py. It was developed in Python and allows for the development of parser-based interactive fiction (known early on as "text adventures"), so that changes in the narrating might be triggered by player input or events in the simulated IF world. Later, Montfort developed Curveship-js in JavaScript. This version of Curveship does not take textual input, as it lacks a parser of the sort used in interactive fiction. It also has a reduced capacity for world simulation. However, it has been extended in certain ways to be a better model of narrative theory and is easy to run in a Web browser. In the current Curveship-js, as much linguistic information as possible is represented separately from the underlying characters, places, things, and events at the story/content level. Curveship-js has been used in teaching about narrative theory as well as for research. Both Curveship-py and Curveship-js are free (libre) software, so all code can be downloaded, studied, and shared.[1] Anyone is allowed to use these systems as the basis of their own research and artistic work — or for any purpose at all.

1 https://nickm.com/curveship/

As an example, here is a Curveship output with a default spin and the naming scheme of a detached narrator:

A bank teller reads a deposit slip.
A burly guard sleeps.
The bank teller rechecks the deposit slip.
A twitchy man puts on a Dora the Explorer mask.
The bank teller types.
She plays Solitaire a bit on her computer.
The twitchy man leaves the street.
The bank teller waves to him.
He threatens her using a gun-shaped object.
She laughs.
The burly guard wakes.
He sees the twitchy man.
He leaves the guard post.
The bank teller puts some fake money into a black bag.
The twitchy man turns to the burly guard.
He shoots him in the chest.
He shoots him in the chest.
He falls.
He dies.
The bank teller weeps.
The end.

Using the same story file to specify events, actors, and things, but given a different spin, Curveship will produce different results. For example, with the bank teller made into the narrator or "I" of the story, with all events elided except those that the bank teller witnesses, and with specific names for things and verbs used to represent events that are appropriate to the bank teller, this is the output:

I glance at a completed Form D-22.
I look over the deposit slip.
I do some data entry.
I play Solitaire.
Jimmy Smith pretends to rob.
I wave to him.
Jimmy poses for me using his gun-shaped object.
I laugh.
Our guard pops out of the guard post.

I place some fake money into a black bag.
Jimmy turns to the guard.
The guard shoots him.
He executes Jimmy.
Jimmy falls.
He dies.
The guard drops his pistol.
He recalls that he shot Jimmy.
I weep.
The guard stares at the pistol.
The end.

5. *Pipelined Integration and the Blackboard of Slant*

We undertook two projects to connect MEXICA and Curveship. Just as developing a single computational model provides insights into the theory or account being modeled, there are benefits to integrating two or more models. We are able to see in concrete terms where representations are compatible or need to be altered, and where basic assumptions differ. For instance, two models can both be formal systems but have representations that are at different levels of granularity. One can include elements that are judged essential while the other omits these.

Initially, we devised a *pipelined* architecture in which MEXICA generates a plot and, given a narrative specification, Curveship determines the particular way the text of the narrative is generated (Montfort & Pérez y Pérez 2008). This simple model seems to us to be more relevant to industrial production than to creative ideation in many ways, and we do not suggest that it is a good model of individual or collaborative creativity. However, it still presented some interesting challenges, because we needed to formulate a compatible representation of story and determine what narrative specification might be reasonable. MEXICA, for instance, had a very different and higher-level representation of action and lacked a representation of particular "props," or things. Curveship was able to deal with finer-grained representation of actions but could not do anything with MEXICA's information about emotional connections between characters. So, we learned specific things about how our models of storytelling were compatible (we did achieve an integration between the models) as well as how they differed.

We extended this connection between MEXICA and Curveship in two ways. First, we introduced a *blackboard* architecture (Hayes-Roth 1985). The blackboard architecture was introduced as a problem-solving technique in AI, but was inspired by human collaboration and, in our formulation of it, is consistent with the E-R Cognitive Account. This scheme does not simply forward partially-complete writing along a pipeline or assembly line. It allows each system component to write in a shared space and collaborate in a richer way. In general, the blackboard can have edits and elements can be erased, although in our system, Slant, we initially only allowed components to augment the representation. To allow the different components to interact, we went around in a sort of circle, giving each component the opportunity to update the blackboard. This represents an enrichment of the E-R cycle.

Thanks to collaborator D. Fox Harrell, we incorporated another component, GRIOT-Gen, that developed figurative conceptual blends. Instead of choosing the narrative specification ourselves, manually, we developed Verso, a fourth component that, in collaboration with the other systems, selected which fine-grained genre would be used to narrate the story (Montfort et al. 2013). Among other insights, we showed that a representation can be shared among a larger number of components with different emphases and that a variety of different contributions can be incorporated into this computer creative writing process.

The following are two example outputs from Slant, with two different plots, two fine-grained genres, two different spin settings, and figuration from GRIOT-Gen applied in the second case. Both examples are from (Montfort et al. 2013):

> Forgive me, Father, for I have sinned. It has been a month since my last confession. An enemy slid. The enemy fell. The enemy injured himself. I located a curative plant. I cured the enemy with the curative plant. The tlatoani kidnapped me. The enemy sought the tlatoani. The enemy travelled. The enemy, um, looked. The enemy found the tlatoani. The enemy observed, uh, the tlatoani. The enemy drew a weapon. The enemy attacked the tlatoani. The enemy killed the tlatoani with a dagger. The enemy rescued me. The enemy entranced, uh, me. I became jealous of the enemy. I killed the enemy with the dagger. I killed myself, uh, with the dagger. Ten Hail Marys? Thank you, Father.

> This is Ehecatl, live from the scene. The cold-wind eagle knight is despising the icy jaguar knight! The cold-wind jaguar knight is despising the chilling eagle knight! Yes, an eagle knight is fighting a jaguar knight! Look at this, the eagle knight is drawing a weapon! Look at this, the eagle knight is closing on the jaguar knight! The gardener eagle knight is wounding the weed jaguar knight! And now, the jaguar knight is bleeding! Yes, the consumed eagle-knight is panicking! And, eagle knight is hiding! Holy — the snowflake slave is despising the chilling jaguar knight! The freezing-wind jaguar knight is despising the cold slave! And, yes, the cold-wind slave is detesting the chilling jaguar knight! A slave is curing the jaguar knight! And, the slave is returning to the city! And, the jaguar knight is suffering! The frozen jaguar knight is dying! Back to you!

In our further discussion we will consider what we learned from MEXICA and Curveship individually. As Slant and the previous integration project shows, however, models of storytelling that we devise don't have to remain separate. Integrating models that work at different levels can enable new sorts of inquiry.

6. *Humanistic Insights*

6.1. Learning from MEXICA

This difference between a cognitive account (as with Sharples's) and a particular computational model (Pérez y Pérez's MEXICA) explains why the construction of a computational model results in a detailed description of the processes and knowledge structures involved when generating narratives. Some details had not been worked out before, but some higher-level processes were also not specified. Implementing these offers new ways of understanding narrative. Here are some examples.

One of the main contributions of MEXICA is the representation of knowledge in terms of emotional relationships and conflicts between characters. There are many plot generation systems that incorporate emotion in some way. For example, TALE-SPIN (Meehan 1976) and MINSTREL (Turner 1993) use variables that characterize the emotional states of characters, and those variables are used as part of the conditions necessary to activate some goals. DAYDREAMER (Mueller 1987) goes further, by employing these types of variables to control the flow of the program, that is, to activate and deactivate goals during the execution of the program. However, MEXICA made a contribution we believe is unique. We do not know of any other system that works with emotional relationships and conflict between characters as a mechanism to progress a story action by action.

MEXICA illustrated how its representation allows sequences of actions to be threaded together in a coherent way. This result is significant since, until then, story generating systems used narrative structures predefined by their designers to ensure coherence. That is, the structure of the story was defined before the program generated it. MEXICA showed the need to expand research on the role of emotions and conflicts as a mechanism to progress toward a finished composition.

Problem solving has traditionally been represented as carrying out a series of actions to achieve a goal. In areas such as cognitive science or artificial intelligence, problem solving has been the basis for characterizing various cognitive processes. This approach has been used in the generation of texts. For example, a knight has the goal of rescuing a princess, so the actions carried out by the knight to achieve the rescue make up the plot. For many years, mainly during the 1980s, 1990s and the beginning of the 21st century, the vast majority of story generating systems used goal representation as the method to develop a piece. MEXICA demonstrates that other mechanisms can guide the generation of narratives. In particular, during Engagement the system produces sequences of actions without using any type of characters goals.

MEXICA-impro (Pérez y Pérez 2015) is a system for the collaborative generation of plots. In it, two MEXICA agents, one called the leader, the other called the follower, work as a team to generate a plot. An important feature of MEXICA-impro is that each of these agents has its own database, which stores knowledge that the program uses to produce narratives. Every time the system generates a new plot, it is added to the database, thus increasing the knowledge of the agent. The generation process works as follows. The leader begins and after one ER cycle sends the follower the material it has generated up to that point. The follower then continues to develop the plot and, after executing one ER cycle, returns the updated version of the text to the leader. Now the leader is the one who continues, after an ER cycle, sends the elaborated plot back to the follower. This back and forth is repeated until the leader decides that the plot is finished.

When the agents' databases are very similar, communication between them flows smoothly, but the plots generated are not very original and contribute little to increase the agents' knowledge. On the other hand, when the databases are very different, communication between the agents is complicated since it is difficult for them to be able to continue what the other generated. If despite these difficulties they manage to produce a finished plot, it is usually what could be called "too original" and can be hard to interpret as coherent and meaningful. In other words, the knowledge it provides is very different from what already exists in the database. But acquiring knowledge that is almost completely disassociated from existing knowledge is of little use in producing new plot. The best results are obtained when the databases are only somewhat similar. This allows reasonably fluid communication, which produces plots novel enough to expand the database, but at the same time generates new knowledge that can be associated with other elements of the database. The outcome of this research provides evidence that it is important in the writing process to balance between having some overlap in knowledge, but significant differences as well, potentially helping us understand collaboration between human writers.

6.2. Learning from Curveship

We'll turn to a quite specific theoretical insight offered by Curveship, an insight about Genette's (1983) concept of *distance,* which he indicates is a type of narrative voice. Distance is "one of two major factors regulating narrative information … The more covert the narratorial mediation and the more numerous the details provided about the narrated situations and events, the smaller the distance that is said to obtain between them and their narration" (Prince 2003). While Genette writes about distance as if it were distinct from speed, time of narrating, and other aspects, Curveship is able to narrate in a way that seems more or less distant simply by varying other aspects of narrative, including these. So, the system provides evidence that distance is a composite of other aspects of narrative rather than its own simple aspect (Montfort 2011). This is of course a somewhat fine-grained insight, not intended to overturn the idea of distance as a useful concept within narrative theory. The way developments in all sorts of theoretical work proceed is often by refinement rather than revolution.

Genette (1983) also proposes a representation of underlying events in the telling that is a numerical sequence: If we have seven underlying events, numbered 1 through 7, our narration can order them 6234517, for instance, or (omitting some by using ellipsis)

RCL — Revista de Comunicação e Linguagens *Journal of Communication and Languages* N.58 (2023) ISSN 2183-7198

1267. Work on designing and implementing Curveship offered another fine-grained insight, showing that a different, richer representation is even more compatible with Genette's narratology and offers increased understanding with only slightly more complexity. Specifically, the way underlying events are presented in the narrative is better represented in a tree structure, with the "now" of narrative at the top level and nodes for flashbacks, flashforwards, and groupings by topic (syllepsis). These nodes can be embedded at lower levels to represent situations in which the narration is elaborate. Different tree representations, which make the different types of narration clear, might be collapsed into a single identical representation, which is less expressive and allows for less understanding, if simply presented as a numerical sequence (Montfort 2011).

Aside from how incredible or ordinary certain plot-level events might be to a reader, a narrator can express surprise, or lack of surprise, at any point in a narrative. After determining that generating expressions of surprise requires a model of both cultural expectation and how easily surprised (or jaded) an individual narrator is, Curveship was used to augment narratives with particular markers of surprise and lack of surprise. In the process, some insights from sociolinguistics were combined with those from narrative theory to determine particular ways that these sorts of expressions could be formally modeled and produced. While many initial insights were generated, the work also exposed the complexity of surprise and lack of surprise expressions (Montfort et al. 2014). Generating them automatically will likely require rich world models operating at four levels: The particular narrator, the particular narrative, the genre of the work (e.g., magical realism or nautical fiction), and more general cultural norms.

A recent Curveship project involved developing more extensive support for referring expressions, noun phrases or surrogates for such phrases (pronouns, restrictive relative clauses, etc.) that represent objects. Curveship's particular focus is on the noun phrases used to indicate characters and objects in the story world. With new support for referring expressions added, and an ability to change the verbs used to represent actions, it has also been possible to explore to what extent the style of particular authors, and even specific books, can be imitated by using referring expressions that seemed suitable. Using identical story files, Curveship was able to generate narrations in the style of some specific works of American and English literature. For instance:

> The type of guy who can get a reservation at Le Bernardin walks to first class. He sits in seat 1B. I notice coolly. A male flight attendant glances at the famous guy's Air Jordan 4 Retro Kaws purchased from Flight Club. He sneers "sir, I believe you're in the wrong cabin" to the famous guy. The famous guy gets his boarding pass from his bespoke Michael Andrews sportscoat. He shows his boarding pass to the male flight attendant. The male flight attendant mumbles "oh, I'm sorry" to him. He pulls out a BIC pen from K-Mart on Astor Place and a Mead memo pad bought at Key Foods. He says "I shouldn't ask, but ... my daughter would really love to have your autograph."

This uniformed devotchka gasped. Sir Harry Styles had held his boarding pass against a scanner. He placed his boarding pass in his carman. He walked to the first class cabin. He sat in seat 1B. This forella reacted. A veck viddied Sir Styles's sabogs. He sneered "sir, I believe you're in the wrong cabin" to Sir Styles. Sir Styles got his boarding pass from his carman. He showed his boarding pass to the veck. The veck muttered "oh, I'm sorry" to him. He grabbed a pen and a notepad. He skazated "I shouldn't ask, but … my daughter would really love to have your autograph" to Sir Styles.

Montfort and collaborators deemed these two results, which were attempts to imitate the style of Bret Easton Ellis's *American Psycho* and Anthony Burgess's *Clockwork Orange,* to be most successful. These are distinctive because of their use of brand names and locations of purchase, in the former case, and an invented dialect with new lexical items (but not much of a change in syntax) in the latter. Using techniques such as the addition of courtesy titles, Curveship also managed to suggest the style of Jane Austen, but less strongly. Austen's writing contains a good deal of commentary and often mentions family and social relations in a way that is harder to model. Discussion of this, along with source code from the story file and one of the narrator files, is provided in (Montfort et al. 2021).

7. *Narrative Models and Large Language Models*

Those who have heard anything recently about natural language processing in general, or story generation in particular, must have heard about ChatGPT (which originally employed "GPT-3.5") and GPT-4, created by OpenAI. ChatGPT and GPT-4, like predecessors GPT-3 and GPT-2, employs a technique known as deep neural networks (DNN) or deep learning, and specifically an autoregressive model developed by researchers at Google less than six years ago, called transformer (Vaswani et al. 2017). The original article on transformer has, as of this writing, already been cited more than 60,000 times. Since we first submitted the paper you are reading now, GPT-4 has been released along with a technical report on this new model (OpenAI 2023).

The "GPT" in OpenAI's systems stands for Generative Pre-Trained Transformer. In pre-training, a system of this sort ingests billions of words — the CommonCrawl data set of Web pages, with 410 billion words, is typically used, along with several smaller data sets. A Large Language Model's neural network has a massive number of different weights, settings, or, more commonly, *parameters.* GPT-3 has 175 billion. That model is also distinguished by its context window of 2049 tokens, meaning that it can consider a very long sequence at once. In the case of ChatGPT, other machine learning methods were used to develop the system, including supervised methods that were employed to keep the model from producing offensive outputs.

The context window for GPT-4 model is 32768 tokens, which easily spans all of Hemingway's *The Old Man and the Sea.* Although OpenAI released a lengthy technical report promoting the system (OpenAI 2023), they documented even fewer details about the system than with previous models. OpenAI's GPT systems are in no way open. They are proprietary trade secrets, exclusively licensed to Microsoft. No peer-reviewed technical paper has been published about GPT-3 or GPT-4.

Thanks to publicity efforts and the clever, selective provision of access to recent GPT models, these have been the topic of a great deal of discussion on social media and in the news, where people have expressed concern about how such new models might automate existing jobs or even lead to artificial general intelligence.

Excitement, adulation, and fear have been directed largely at OpenAI's models, and to some extent at Google's Bard, but there are a wide variety of LLMs. The BigScience Large Open-science Open-access Multilingual Language Model (BLOOM) is, as the name says, a free and open model and tops GPT-3 with 176 billion parameters. Google Research has developed several proprietary models, including the Pathways Language Model (PaLM) with 540 billion parameters and Generalist Language Model (GLaM) with 1 trillion parameters.

While deep neural nets, autoregressive learning, and the transformer technique are not easily described, we can mention a bit about the essential idea behind how a Large Language Model operates. For instance, we can compare an LLM to a very simple type of generative model, the Markov chain or Markov process. A typical way in which this process would be implemented and used, for instance by a computer science student today undertaking an assignment, would be to have a program read in a book's worth of text and generate new textual productions that look one word behind the current word to consider what text to produce next. GPT-3 in contrast is able to consider not just the previous word, but four single-spaced pages worth of context, and rather than having read in a single book, it has read the equivalent of immense libraries, thanks to text available on the internet. GPT-4 goes far beyond that in context. The deep neural network architecture of an LLM also means that it can consider a sequence of words that is entirely novel and determine a probability distribution of words to follow it, something no Markov chain could do. Because LLMs can consider extremely long and novel word sequences and determine which words are most likely to follow, they are able to accomplish unprecedented and uncanny continuations of existing text.

LLMs are certainly formidable when compared to Markov chain generators, although at a very high level, the idea is similar. They predict, given some amount of language, what textual output is likely to follow the text provided. As we look closer at these systems and what exactly they model, we'll turn to the latest and greatest system to which we had access when we wrote this article: ChatGPT.

As you might expect from this comparison, the texts output by ChatGPT and GPT-4 exhibit impressive cohesion. One of the other remarkable characteristics of these systems, due to the immense amount of training they have received, is that they can produce replies, including narrative replies, about an enormous number of topics, seeming to function as competent storytellers. There are also numerous limitations of these models — their replies are often incorrect and, without special work from system designers to censor them, can be offensive and unethical — but to conclude our discussion here, we will focus on just one limitation.

Fundamentally, LLMs, including GPT-3 and the new and improved ChatGPT and GPT-4, are simply not models of *writing* (how human beings write) or of *narrative* (how human beings tell stories). A language model is simply a different sort of model, which can accomplish different tasks but should not be expected to offer insights into writing and narrative in the same way MEXICA and Curveship have.

Let us elaborate this idea. LLMs "know" how to continue word sequences. Given the text they are trained on, they produce an amazingly good probability distribution of words that would follow sequences of words, including very long sequences, including sequences that do not occur anywhere in the training data. Such probability distributions can be very useful in speech recognition and machine translation, for instance. But these models do not "know" anything about plots, about narrative, or even about grammar (Chomsky et al. 2023). They just "know" what word is likely to come next.

Because an LLM is a model, it has external parameters, which include the *temperature.* A low temperature gives very conservative and ordinary words. Indeed, if the temperature is turned all the way down to 0, there will be a deterministic result. A high temperature provides more unusual and unconventional outputs. Another parameter is *top-k,* which limits the possible choices to a list of k different ones. Like other LLM parameters, these pertain to the statistics of the text the model was trained on.

In contrast, MEXICA is fundamentally agnostic to natural language, which is why its plots can be rendered at the end of the process into English as easily as Spanish. MEXICA as a model embodies ideas about how to resolve tension (for example), but its main function is not modeling how to express the plot it generates. Some of the system parameters allow deciding the pace of development of a narrative and determining how similar the story context should be to the knowledge structures in the database.

Curveship, by contrast, is a text generator, but does not function by determining an appropriate next word. It takes an abstract representation of a narrative and renders that in a particular narrative style. Curveship's parameters are, collectively, called *spin,* as in the spin that one puts on a story. One of the parameters controls who the "I" or narrator of the story is — which character, if any, will be in that role. Another gives the ordering of events, which might result in events being told out of chronological order but might also cause ellipsis, the omission of certain events. The plot in every case remains the same: Curveship's parameters are all about how the plot is actually expressed or told.

We hope this discussion explains why LLMs, however impressive, don't do the same thing as models of human writing processes or human narrating. It's also worth noting that the particular proprietary system we used here, ChatGPT, has already been changed and tuned up by OpenAI, making the closed "December 15" version that we used inaccessible. This means it is impossible for researchers in the future to repeat our experiment. OpenAI (a corporation dedicated, again, to closed systems) is bent on automating customer service representatives and providing entertaining chatter, and has no stake in advancing humanistic knowledge and understanding. However, there are LLMs which, like Curveship for instance, are free and open source. These include BLOOM, the English-based GPT-NeoX with 20 billion parameters and another model, pre-trained with French-language data and using 6 billion parameters, called Boris. We expect these models may be able to allow us to learn, for instance, about the relationship between English and French as represented in digital and digitalized writing. Even then, we should be aware that this is not the purpose for which LLMs are created. In this way, they differ from research models such as MEXICA and Curveship which are specifically made for purposes of particular types of inquiry.

Conclusion

ChatGPT and GPT-4 are both closed, proprietary models; they change all the time so experiments are not reproducible; and the actual numerical parameters of the LLM are about probability distributions of words, not plot or narrative style. By contrast, we have built models of plot and narrative for the specific purpose of inquiring about these, not about the way words follow from one another. The construction of MEXICA, Curveship, and systems like them follow four main steps:

1. Development of a cognitive/narrative model that describes aspects of narrative generation.
2. Transformation of that cognitive/narrative model into a computer model.
3. A detailed study of how each of the elements in the computer model interact, and how they manipulate and transform information.
4. Evaluation of the outputs produced by the systems and analysis of the relation between the output's features and the elements and parameters of the model.

Each step in their construction can contribute to our overall understanding of the generation of narratives.

Stories are essential to human beings and our communication, the ways we process experiences, and how we make sense of the world. However, we still know too little about the mechanisms necessary to create and tell stories. Some systems, including MEXICA, computationally model the cognitive processes associated with creative writing. Others, including Curveship, computationally represent methods of narrating. These systems are also open and allow for parametric changes and experimentation. However impressive and surprising recent large language models may be, they do not serve the purposes of models such as MEXICA and Curveship. The types of inquiry we can undertake with these domain-specific research systems give us powerful new ways to learn about human storytelling and possibilities for narration, building on cognitive accounts and humanistic theories.

Bibliography

Alhussain, Arwa I., and Aqil M. Azmi. 2021. "Automatic Story Generation: A Survey of Approaches." *ACM Computing Surveys (CSUR)* 54(5).1-38. https://doi.org/10.1145/3453156.

Chomsky, Noam, Ian Roberts, and Jeffrey Watumull. 2023. "The False Promise of ChatGPT." *The New York Times,* March 8, 2023.

Genette, Gérard. 1983. *Narrative Discourse: An Essay in Method*. Ithaca. NY: Cornell University Press.

____________. 1988. *Narrative Discourse Revisited*. Ithaca, NY: Cornell University Press.

Gervás, Pablo. 2009. "Computational Approaches to Storytelling and Creativity." *AI Magazine* 30(3): 49-62. https://doi.org/10.1609/aimag.v30i3.2250.

Hayes-Roth, Barbara. 1985. "A Blackboard Architecture for Control." *Artificial intelligence* 26(3):251-321. https://doi.org/10.1016/0004-3702(85)90063-3.

Herrera-González, Brian Daniel, Alexander Gelbukh, and Hiram Calvo. "Automatic Story Generation: State of the Art and Recent Trends." In *Advances in Computational Intelligence. MICAI 2020. Lecture Notes in Computer Science*, vol 12469, edited by Lourdes Martínez-Villaseñor, Oscar Herrera-Alcántara, Hiram Ponce, and Félix-A. Castro-Espinoza, 81-91. Springer International Publishing. https://doi.org/10.1007/978-3-030-60887-3_8.

Hou, Chenglong, Chensong Zhou, Kun Zhou, Jinan Sun, and Sisi Xuanyuan. 2019. "A Survey of Deep Learning Applied to Story Generation." In *Smart Computing and Communication. SmartCom 2019. Lecture Notes in Computer Science()*, edited by Meikang Qiu, vol 11910, 1-10. Springer. https://doi.org/10.1007/978-3-030-34139-8_1.

Kybartas, Ben, and Rafael Bidarra. 2016. "A Survey on Story Generation Techniques for Authoring Computational Narratives." *IEEE Transactions on Computational Intelligence and AI in Games,* 9(3): 239-253. https://doi.org/10.1109/TCIAIG.2016.2546063.

Meehan, James R. 1976. "The Metanovel: Writing Stories by Computer." PhD Thesis Dissertation, Yale University. https://apps.dtic.mil/sti/citations/ADA031625.

Montfort, Nick, Rafael Pérez y Pérez, D. Fox Harrell, and Andrew Campana. 2013. "Slant: A Blackboard System to Generate Plot, Figuration, and Narrative Discourse Aspects of Stories." *Proceedings of the International Conference on Computational Creativity (ICCC)*. https://hdl.handle.net/1721.1/137198.2.

Montfort, Nick. 2007. "Generating Narrative Variation in Interactive Fiction." PhD dissertation, University of Pennsylvania. https://repository.upenn.edu/dissertations/AAI3271851.

____________. 2011. "Curveship's Automatic Narrative Style." In *Proceedings of the 6th International Conference on the Foundations of Digital Games* (FDG '11), edited by Georgios N. Yannakakis, Antonios Liapis, Penny Kyburz, Vanessa Volz, Foaad Khosmood, and Phil Lopes, 211-218. New York, United States: Association for Computing Machinery. http://hdl.handle.net/1721.1/67645.

Montfort, Nick, and Rafael Pérez y Pérez. 2008. "Integrating a Plot Generator and an Automatic Narrator to Create and Tell Stories." In *Proceedings of the 5th International Joint Workshop on Computational Creativity,* edited by Pablo Gervás, Rafael Pérez y Pérez and Tony Veale, 61-70. Madrid: Universidad Complutense de Madrid. https://computationalcreativity.net/ijwcc08/proceedings.html.

Montfort, Nick, Ardalan SadeghiKivi, Joanne Yuan, and Alan Y. Zhu. 2011. "Using Referring Expression Generation to Model Literary Style." In *Proceedings of the Workshop on Natural Language Processing for Digital Humanities (NLP4DH),* edited by Mika Hämäläinen, Khalid Alnajjar, Niko Partanen, and Jack Rueter, 64–74. Taiwan, Taipei: Association for Computational Linguistics. https://aclanthology.org/2022.nlp4dh-1.0.pdf.

Montfort, Nick, Erik Stayton, and Andrew Campana. 2014. "Expressing the Narrator's Expectations." *Intelligent Narrative Technologies 7 (INT7),* 24–30.

Mueller, Eric T. 1990. *Daydreaming in Humans and Machines: A Computer Model of the Stream of Thought.* Norwood, NJ: Ablex Publishing Corporation.

Open AI. 2023. *GPT-4 Technical Report*. Version 2. San Francisco: OpenAI.. https://doi.org/10.48550/arXiv.2303.08774.
Pérez y Pérez, Rafael. 1999. "MEXICA: A Computer Model of Creativity in Writing." DPhil Dissertation, University of Sussex. http://www.rafaelperezyperez.com/documents/MEXICA_DPhil_RPyP.pdf.
____________. 2007. "Employing Emotions to Drive Plot Generation in a Computer-Based Storyteller." *Cognitive Systems Research* 8(2):89-109. https://doi.org/10.1016/j.cogsys.2006.10.001.
____________. 2015. "A Computer-based Model for Collaborative Narrative Generation." *Cognitive Systems Research*, 36-37: 30-48. https://doi.org/10.1016/j.cogsys.2015.06.002.
____________. 2017. *MEXICA: 20 Years, 20 Stories / 20 Años, 20 Historias*. Using Electricity series. Denver: Counterpath.
____________. 2018. "Comunicación e Interdisciplina: Perspectiva desde la Creatividad Computacional." En *Sociedad del conocimiento y comunicación: reflexiones críticas desde América Latina,* coordenado por Delia Crovi Druetta,199-210. Ciudad de México: La Biblioteca.
Pérez y Pérez, Rafael, and Mike Sharples. 2001. "MEXICA: A Computer Model of a Cognitive Account of Creative Writing." *Journal of Experimental and Theoretical Artificial Intelligence* 13(2):119-139. https://doi.org/10.1080/09528130010029820.
____________. 2004. "Three Computer-Based Models of Storytelling: BRUTUS, MINSTREL and MEXICA." *Knowledge Based Systems Journal* 17(1): 15-29. https://doi.org/10.1016/S0950-7051(03)00048-0.
Pérez y Pérez, Rafael, and Mike Sharples. *An Introduction to Narrative Generators: How Computers Create Works of Fiction*. Oxford University Press, Forthcoming.
Prince, Gerald. 1982. *Narratology: The Form and Functioning of Narrative*. Berlin: De Gruyter Mouton.
____________. 2003. *A Dictionary of Narratology*. Revised edition. Lincoln: University of Nebraska.
Queneau, Raymond. 1981. *Exercises in Style*. Translated by Barbara Wright. New York: New Directions.
Reiter, Ehud, and Robert Dale. 2000. *Building Natural Language Generation Systems*. Studies in Natural Language Processing series. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519857.
Ryan, Marie-Laure. 1991. *Possible Worlds, Artificial Intelligence, and Narrative Theory*. Bloomington: Indiana University Press.
____________. 2001. *Narrative as Virtual Reality*. Baltimore: Johns Hopkins University Press.
Schank, Roger C. 1972. "Conceptual Dependency: A Theory of Natural Language Understanding." *Cognitive Psychology* 3(4): 552-631. https://doi.org/10.1016/0010-0285(72)90022-9.
Sharples, Mike. 1999. *How We Write: Writing as Creative Design*. London: Routledge.
Sharples, Mike, and Rafael Pérez y Pérez. 2022. *Story Machines: How Computers Have Become Creative Writers*. London and New York: Routledge, Taylor and Francis Group.
Turner Scott R. 1994. *The Creative Process: A Computer Model of Storytelling*. Hillsdale NJ: Lawrence Erlbaum Associates.
Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. "Attention is All You Need." *Advances in Neural Information Processing Systems* 30, edited by I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett. Long Beach, California, USA. https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

—

Funding

This research was partially financed by the Research Council of Norway through its Centres of Excellence Scheme, Project No. 332643.

NICK MONTFORT

Nick Montfort is a poet and artist who uses computation as his main medium and seeks to uncover how computing and language are entangled with each other and with culture. His computer-generated books include *#!* and *Golem.* His digital projects include the collaborations *The Deletionist* and *Sea and Spar Between.* Montfort also studies creative computing. MIT Press has published his *The New Media Reader, Twisty Little Passages, The Future,* and *Exploratory Programming for the Arts and Humanities.* He directs a lab/studio, The Trope Tank, and is professor of digital media at MIT. He lives in New York City.

—
ORCID
0000-0001-7558-5160
—
Institutional address
Massachusetts Institute of Technology
77 Massachusetts Ave, 14E-316
Cambridge, MA 02139
United States

RAFAEL PÉREZ Y PÉREZ

Rafael Pérez y Pérez is a Research Professor at Universidad Autónoma Metropolitana at Cuajimalpa, México City. In 2006 he founded the Interdisciplinary Group on Computational Creativity, which aims to gather together a group of researchers and students interested in computational creativity. They have developed programs for plot generation, interior design, visual narratives, creative problem solving, and so on. This group organizes every year the Mexican International Colloquium on Computational Creativity. He was the chair (2014-2015, 2015-2019) of the Association for Computational Creativity. Professor Pérez y Pérez is member of the National System of Researchers in México (SNI).

—
ORCID
0000-0001-8985-5897
—
Institutional address
Universidad Autónoma Metropolitana
División de Ciencias de la Comunicación y Diseño.
Av. Vasco de Quiroga 4871, Col. Santa Fe
Cuajimalpa, C.P. 05348.
Ciudad de México, México.

—
To cite this article
Montfort, Nick, and Rafael Pérez y Pérez. 2023. "Computational Models for Understanding Narrative." *Revista de Comunicação e Linguagens* (58): 97-117. https://doi.org/10.34619/gnzq-r7ri.

—
Received Recebido: 2022-12-24 **Accepted** Aceite: 2023-04-24

—
DOI https://doi.org/10.34619/gnzq-r7ri