Unmet Creativity Support Needs in Computationally Supported Creative Writing

Max Kreminski
University of California, Santa Cruz
mkremins@ucsc.edu

Chris Martens
North Carolina State University
martens@csc.ncsu.edu

Abstract

Large language models (LLMs) enabled by the datasets and computing power of the last decade have recently gained popularity for their capacity to generate plausible natural language text from human-provided prompts. This ability makes them appealing to fiction writers as prospective co-creative agents, addressing the common challenge of writer's block, or *getting unstuck*. However, creative writers face additional challenges, including maintaining narrative consistency, developing plot structure, architecting reader experience, and refining their expressive intent, which are not well-addressed by current LLM-backed tools. In this paper, we define these needs by grounding them in cognitive and theoretical literature, then survey previous computational narrative research that holds promise for supporting each of them in a co-creative setting.

1 Introduction

Mixed-initiative co-creative (Liapis et al., 2016; Deterding et al., 2017) creativity support tools (Shneiderman, 2007) for creative writing have recently seen a surge of interest in research communities, coinciding with the introduction of large language models (LLMs) such as GPT-3 (Brown et al., 2020) that can provide coherent suggestions for the continuation of human-written text. Several recent efforts have been made to understand the experiences of writers who work with these tools to produce texts (Manjavacas et al., 2017; Roemmele and Gordon, 2018; Calderwood et al., 2020). However, less attention has been paid to the development of systems that can provide forms of creative writing support beyond short-term suggestions for textual continuation.

Meanwhile, recent efforts to understand the playful creative writing communities that have emerged around interactive emergent narrative games (Kreminski et al., 2019b; Kreminski and Wardrip-Fruin, 2019) and to provide computational support for playful creative writing at the plot-structure level (Kreminski et al., 2020a) have revealed a preliminary inventory of several distinct but interrelated creativity support needs among creative writers, including:

- Getting unstuck
- Maintaining consistency
- Constructing a satisfying overall story arc, including a conclusion/resolution
- Managing reader experience
- Refining and iterating on expressive intent

Current large language models are good at addressing the first of these needs, *getting unstuck*, via short-term suggestions that can prompt writers to take their stories in unexpected new directions. However, they do not directly address consistency maintenance, longer-term plot structure, management of reader experience, or the challenge of refining high-level expressive intent, and some novelists even suggest that LLMs may actively work against the construction of coherent plot structure due to the highly divergent nature of LLM suggestions (Calderwood et al., 2020). Some recent work aims to improve LLMs in ways that could enable them to meet these needs: for instance, work in long text generation (Hua and Wang, 2020; Guan et al., 2021; Tan et al., 2021) could assist users with consistency maintenance; work on hierarchical concept-driven language models (Wang et al., 2021) could help to maintain plot structure in generated text; and work in diverse decoding methods (Ippolito et al., 2019; See et al., 2019) could help users refine their intent by selecting from among diverse potential completions of the same text. However, the possibility of supporting these needs through other forms of technology may also be worth investigating.

In this paper, we describe each of these creative writing support needs in more detail, then survey previous research from communities outside of NLP/computational linguistics that have either been shown capable of addressing, or that show potential for supporting these creative needs. Our aim with this paper is to create a bridge between the ACL community and AI/digital games research community that may yield productive insight towards synthesizing these approaches that have evolved in parallel.

We limit the scope of our discussion primarily to narrative fiction, particularly in the form of short stories, novels, and game writing/interactive storytelling, so the suggestions made here may not all be applicable to other forms of creative writing (such as poetry). However, we attempt to avoid limiting ourselves to purely text-based storytelling in which only the written word is used to convey meaning; we are also interested in forms of narrative fiction that target visual, audio, and hybrid renderings of fictional events, such as film and game narrative, since many technologies capable of reasoning about plot structure are readily applicable to these domains.

2 Creative Writing Support Needs

2.1 Getting Unstuck

One common source of difficulty in creative writing is the prevalence of *writer's block*, or the sense that one has become "stuck" and cannot think of any obvious way for the story to proceed. Because writer's block is frequently experienced by writers and difficult to escape, it is often discussed in guides for writers, along with descriptions of exercises and practices that can help prevent writers from becoming blocked or enable them to become unblocked (Lamott, 2007). These exercises and practices take many forms, but they often involve the use of genre-typical plot devices to advance the action in lieu of any more natural continuation (e.g., Raymond Chandler's oft-cited description of a genre-typical move in hardboiled detective fiction: "When in doubt have a man come through the door with a gun in his hand" (Chandler, 1950)) and the use of unfiltered stream-of-consciousness writing for a fixed amount of time (e.g., one hour each day) to help writers continue working through a block (Goldberg, 2005).

It is in helping writers get unstuck that the strengths of large language models are especially apparent. Language model continuations of human-written text tend to be syntactically valid and relevant to storyworld entities or situations that were described in the immediately preceding text, enabling them to function as viable short-term suggestions for what might happen next in a written story. This is true even though these suggestions may sometimes take the story in unexpected or unwanted directions: regardless of whether users accept the suggestions that are provided, co-writing with a language model can shift the user's task from the wholesale invention of a new direction for the story to take (the precise thing that it is difficult to do when blocked) toward the acceptance or rejection of computer-provided suggestions. The latter task can be subjectively easier to perform (Smith, 2012, p. 57), and once a desirable continuation is located, further plot events may occur to the user naturally even without ongoing computational support.

2.2 Maintaining Consistency

When constructing a work of fiction, the author aims to convey a mental model of an underlying *story world*: a set of characters, settings, objects, and relationships between all of these things that change over the course of narrative events according to certain logics that may or may not rely on real-world, non-fictional analogs. Practicing novelists often maintain (and advise beginning writers to maintain) "story bibles" or other collections of extradiegetic "storywork" apart from the narrative text itself that serve to document story world information (Ousby, 2009). The use of story world documentation points to a need to *maintain consistency* in works of fiction. As stories and their casts of characters grow in size, and more of the fictional timeline is filled in, the author runs increasing risk of introducing inconsistencies (conflicting factual assertions or implications), plot holes, or unexplained situations that may break the reader's ability to suspend disbelief.

In order to reason about consistency, authors need to reason about narrative material at a level more abstract than narrative text (including storyboards, scene scripts, etc). It can be useful to reason about the story world and its logic—the *represented* phenomena—separately from the story artifact itself—the *representation of* those phenomena. This distinction basically aligns with the classical Russian narratologists' distinction between

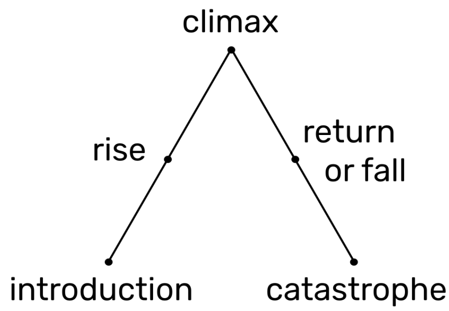

Figure 1: Freytag's pyramid

fabula and *syuzhet* (Gorman, 2018), or its adaptation in anglophone narratology as *story* versus *discourse* (Chatman, 1980). Correspondingly, cognitive linguists have long recognized the presence of *situation models* as knowledge structures that readers create to interpret the semantic relationships between referents in natural language sequences (Zwaan and Radvansky, 1998). The ability to directly author and manipulate knowledge corresponding to a situation model (or similar) is central to a fiction author's task.

2.3 Plot Structure

When writers think about *plot structure*, they may have in mind a set of "acts" (as in "3-act structure") or a continuous curve describing the dramatic tension of the story over time, as in Freytag's pyramid (Freytag, 1894). Although the notion of conflict is not universal (Hunter, 2016), usually, a plot follows a sequence of identifiable *beats* that include establishment of an initial situation, and inciting incident or a need that spurs characters to action, a series of events in which the characters attempt to address the inciting incident, an emotional peak that resolves it, and a denouement or resolution that describes the aftermath (see Figure 1). A number of conceptual models have been proposed and used for describing plot structure, such as the Freytag pyramid, the Monomyth or Hero's Journey (Campbell, 2008), and Dan Harmon's Story Circle (O'Meara, 2015).

Importantly, plot structures describe global rather than local features of a text, and they have more to do with the underlying world model (see previous section) than they do with the specific actions or events that are inferable from lexical properties of the text. Cohn and colleagues have established that readers make sense of stories in a "grammatical" way akin to parsing sentences: they expect certain structures that parse the entire story into something story-like, and in the absence of these structures, comprehension falters (Cohn, 2020).

2.4 Reader Experience

The movement of "human-centered design" proposes that designers benefit when they make an effort to empathize with users: by understanding the experience of the people who will experience and interact with the designed work, we can more intentionally shape those experiences. Likewise, a written work has an experiential impact on its readers, and understanding the levers that affect that impact is a key part of narrative intelligence.

Three examples of reader experience are **pacing**, **tension**, and **surprise**. Pacing refers to the amount of time that a reader spends with each segment, scene, or act of the overall plot (see previous section on plot structure). Poor pacing can cause a reader to get bored or overwhelmed with the story and fail to connect with the characters or the underlying message that the writer is attempting to convey. Tension refers to elements of conflict, threat, or suspense, that cause discomfort in the reader and evoke a sense of wanting the tension to resolve, pushing them forward in the story to feel relief. Surprise refers to encountering unexpected narrative events that shift the reader's mental model of the story and, if done well, increase the reader's curiosity to reconcile their failure to predict what would happen.

Reasoning about reader experience requires a good understanding of how stories work at a cognitive level: e.g., that readers work as *problem solvers* when processing narrative text, working to stay one step ahead of the story to make sense of what has happened so far and predict what will happen next (Gerrig and Bernardo, 1994). If story authors strategically *withhold* information, they can *elicit inferences* on the part of readers to fill in the gaps in ways that can evoke humor, shock, or horror understanding (Cohn, 2019).

2.5 Refining Expressive Intent

One difficulty in creative work is that the creator themselves may not know exactly what they are trying to express, and the expressive intent may shift as the creator's understanding of the work evolves. This is particularly true in storytelling: for instance, a writer's understanding of a particular character's personality may shift (often becoming more nuanced over time) as the writer develops a deeper backstory for the character and places them in plot situations that allow different aspects of the character's personality to come to the forefront. Similarly, the originally intended ending for a story may come to feel inconsistent with the author's better understanding of the story's intended themes partway through the writing process. Divergent suggestions provided by computational support tools may exacerbate these difficulties, making it harder (rather than easier) for writers to "find the heart" of what they are trying to express.

Consequently, it may be helpful for computational support tools to explicitly ask the user about their high-level expressive intent; provide them with a place to write down and edit their intent, perhaps in a machine-understandable form; infer expressive goals from what the user has already written, perhaps allowing them to accept or reject suggestions as to what high-level goals they were trying to accomplish with a particular span of text; and try to provide suggestions that are consistent with the user's high-level expressive goals. Several design patterns for "reflective creators" (Kreminski and Mateas, 2021)—a particular genre of creativity support tools that aim to help users refine their intent—may be of use in this context.

3 Technologies and Approaches

In this section, we overview technologies that have shown promise for addressing the needs outlined in the previous section.

3.1 Maintaining Consistency

The key technological tool for maintaining consistency is a *world model*, or a computational representation of the diegetic phenomena that a story aims to fictionalize. These phenomena include characters (and potentially their interior phenomena such as their personalities and beliefs), settings, character relationships, and narrative actions or events that can modify the world. By representing a world model in its own right, one can specify consistency constraints as (e.g.) first-order logic formulas whose constituent predicates refer to the world model.

World models appear in a number of computational narrative tools. For example, the *stories as plans* approach began as an observation that generating consistent narratives could be cast as an automated planning problem, for which there exist efficient solvers (Young, 1999). Given a description of narrative action schema in terms of their preconditions and effects, and a description of an initial and target story world state, planners generate sequences of narrative actions that are *consistent* in the sense that each action's preconditions are met by the implied world state following the prefix of the sequence leading up to it. Figure 2 shows an example story generation problem set up in this manner, alongside a planner's output. This observation has led to a long history of plan-based approaches to narrative generation (Porteous et al., 2010; Riedl and Young, 2010; Ware and Young, 2011; Young et al., 2013) as well as ongoing research that aims to incorporate more robust models of character intention and belief (Eger and Martens, 2017; Shirvani et al., 2017, 2018; Wadsley and Ryan, 2013).

The *stories as proofs* approach is closely related to planning in that it also relies on a solver to generate logical sequences of events that can be interpreted as consistent stories (Bosser et al., 2010; Martens et al., 2013, 2014); the solver in this case is a linear logic theorem prover (or logic programming language) that can be run in a non-goal-directed (forward chaining) mode, leading to increased solution diversity. The forward-chaining mode also enables a natural introduction of user interaction, allowing a human to "steer" the search process by selecting from among all possible actions (whose preconditions are met in the current world state). This approach suggests opportunities for incorporating world models into a human-centered writing practice, affording levers for authors to express and enforce story consistency.

3.2 Plot Structure

Machine-learned language models are good at capturing local coherence, but tend to struggle with the global constraints implied by plot structure. In direct mappings from text corpora to text output, these structures are at best latent properties of edge weights in a neural network, rather than rules that can be inspected and modified with authorial control.

By contrast, symbolic representation techniques like context-free grammars and logic programming provide a high degree of expressive control. For instance, Gervas (Gervás, 2013) encodes Vladimir Propp's narratological functions as a BNF gram-

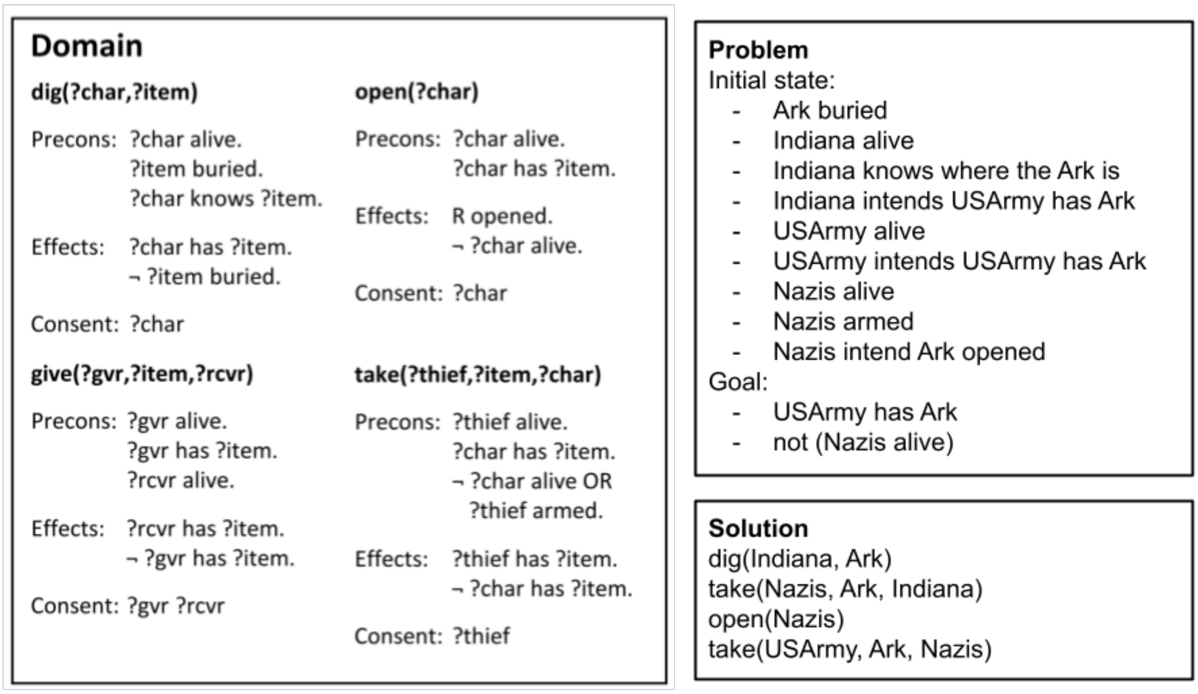

Figure 2: Example planning domain and problem (input) and sample solution plan (output) courtesy of Ware and Young (Ware and Young, 2014).

mar whose expansions correspond to example plots of Russian folktales that Propp's work was designed to describe. Likewise, Cohn's grammar for the visual narrative structure of short comic strips has been implemented as a comic-generating algorithm (Martens and Cardona-Rivera, 2016).

BRUTUS (Bringsjord and Ferrucci, 1999) is an example from the 1990s in which high-level plot structure patterns, such as "one character betrays another," are specified as first-order logic rules that can be written in Prolog and over which queries can be run to generate example narratives that fit a given plot structure. More recently, answer set programming has been used to codify the narrative planning techniques discussed in the previous section, on which plot structure constraints can then be layered (Dabral and Martens, 2020).

3.3 Reader Experience

To support authors in crafting an intentional experience for their readers, computational tools need to be able to reason about (or perhaps even simulate) the reader's cognitive processes. Distinguishing between story and discourse is one promising first step for reader experience support, since it allows a narrative generation engine to retell the same story (plot-wise) in different ways (Rishes et al., 2013). When generating narrative *discourse*, it is possible to relate the told portion of the story to its underlying world model and add a layer of modeling for what the reader (or viewer) will know and infer based on what they have been shown. Jhala and Young's cinematic discourse engine does exactly this in order to plan camera shots for scenes taking place in 3D worlds (Jhala and Young, 2010)

Drama managers are another compelling tool from the interactive storytelling community that bring to bear on reader experience (Roberts and Isbell, 2007). They are conceived as storytelling agents that track player choices throughout the narrative and coordinate the characters and objects in the world to steer the player and the story toward convergent goals. They sometimes generate or select narrative content appropriate to the emergent properties of the situation, as in the breakaway interactive drama Façade (Mateas and Stern, 2003). Such tools could allow authors to tag story content with world model-relevant properties in similar ways, then work with a drama management tool to remix and recombine passages of text as they draft the scene-by-scene structure.

Finally, technologies have been created for modeling reader cognition to support reader experience effects such as pacing, tension, and surprise. The IDTension system uses a world model and the story-discourse distinction to model tension in an interactive drama setting (Szilas, 2003); the Suspenser system models the reader's inference generation process as a planning algorithm (Cheong and Young, 2006). Graesser and Franklin's QUEST model of reader understanding describes the narrative comprehension process as measured through their ability to answer questions, and describes a *knowledge structure* that encodes this question-answering ability (Graesser and Franklin, 1990), and Cardona-Rivera et al. have implemented the QUEST model as an algorithm to annotating generated story content with relevant reader inferences according to this model (Cardona-Rivera and Young, 2019).

3.4 Refining Expressive Intent

Since refinement of expressive intent has only recently been recognized as an explicit goal for creativity support tools in some contexts, relatively little work has been done to provide computational support for intent refinement in storytelling contexts. However, Writing Buddy (Samuel et al., 2016), Mimisbrunnur (Stefnisson and Thue, 2018), and *Why Are We Like This?* (Kreminski et al., 2020a,b) all address this challenge to some extent by providing explicit interfaces for the specification of *author goals*: high-level, machine-interpretable descriptions of what the human user wants to have happen in the story they are writing. These systems then use this information to provide suggestions for story events or storyworld state updates that respect the user's goals, simultaneously assisting users in reflecting on their own goals (by asking them to state these goals explicitly) and in maintaining consistency with these goals (by using goal descriptions to steer suggestions).

Additionally, *story sifting* technologies (Ryan et al., 2015; Ryan, 2018; Kreminski et al., 2019a)—which apply pattern matching to the identification of potentially compelling new plot directions in chronicles of past story events—can also be applied to the task of inferring an author's intent for the story they are writing. If an intelligent writing tool can use story sifting to discover the beginnings of a potentially interesting plot thread are discovered via story sifting, it can then explicitly ask the user whether the narrative direction implied by this plot thread is of interest to them; regardless of the user's answer, this information can be used to interactively build up an explicit model of what the user does and does not want to happen within the story they are telling.

4 Conclusion

We have presented five creative writing support needs, only one of which (getting unstuck) is meaningfully supported by current large language models, and surveyed technologies for addressing the remaining four needs that have arisen from the AI/digital games research community. These technologies are at varying levels of maturity, and most of them have only been tested in purely automated or generative forms rather than in mixed-initiative, co-creative interaction modes. An important line of future work will be to evaluate these technologies in those modes and determine interfaces and interaction protocols that amplify and foster human creativity in the writing process.

Our goal with this paper is not to assert the superiority of world-model or knowledge-engineering based approaches over LLMs, but rather to emphasize that there is a set of needs and affordances that these techniques can address and provide that are complementary to the needs addressed and affordances provided by LLMs. By bridging research communities focused (on one hand) on computing with natural language and (on the other) on simulating story worlds and reasoning about narrative structure, we hope to pave the way for hybrid and unified models that can transform the human creative writing experience—much like the neurosymbolic approaches to automated story generation (Martin, 2021) that undergird several recent advances in story generation as a field.

References

Anne-Gwenn Bosser, Marc O Cavazza, and Ronan Champagnat. 2010. Linear logic for non-linear storytelling. In *19th European Conference on Artificial Intelligence*, pages 713–718. IOS Press.

Selmer Bringsjord and David Ferrucci. 1999. *Artificial intelligence and literary creativity: Inside the mind of BRUTUS, a storytelling machine*. Psychology Press.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901.

Alex Calderwood, Vivian Qiu, Katy Ilonka Gero, and Lydia B Chilton. 2020. How novelists use generative language models: An exploratory user study. In *HAI-GEN + user2agent @ IUI*.

Joseph Campbell. 2008. *The Hero with a Thousand Faces*, volume 17. New World Library.

Rogelio E. Cardona-Rivera and R. Michael Young. 2019. Desiderata for a computational model of human online narrative sensemaking. In *Working Notes of the 2019 AAAI Spring Symposium on Story-enabled Intelligence*.

Raymond Chandler. 1950. The simple art of murder. *Saturday Review of Literature*.

Seymour B. Chatman. 1980. *Story and Discourse: Narrative Structure in Fiction and Film*. Cornell University Press.

Yun-Gyung Cheong and R Michael Young. 2006. A computational model of narrative generation for suspense. In *AAAI*, pages 1906–1907.

Neil Cohn. 2019. Being explicit about the implicit: inference generating techniques in visual narrative. *Language and Cognition*, 11(1):66–97.

Neil Cohn. 2020. Your brain on comics: A cognitive model of visual narrative comprehension. *Topics in cognitive science*, 12(1):352–386.

Chinmaya Dabral and Chris Martens. 2020. Generating explorable narrative spaces with answer set programming. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 16, pages 45–51.

Sebastian Deterding, Jonathan Hook, Rebecca Fiebrink, Marco Gillies, Jeremy Gow, Memo Akten, Gillian Smith, Antonios Liapis, and Kate Compton. 2017. Mixed-initiative creative interfaces. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 628–635.

Markus Eger and Chris Martens. 2017. Character beliefs in story generation. In *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Gustav Freytag. 1894. *Freytag's Technique of the Drama*. Scott, Foresman.

Richard J. Gerrig and Allan B. I. Bernardo. 1994. Readers as problem-solvers in the experience of suspense. *Poetics*, 22(6):459–472.

Pablo Gervás. 2013. Propp's morphology of the folk tale as a grammar for generation. In *2013 Workshop on Computational Models of Narrative*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Natalie Goldberg. 2005. *Writing Down the Bones: Freeing the Writer Within*. Shambhala.

David Gorman. 2018. Russian formalism. *A Companion to Literary Theory*, pages 36–47.

Arthur C Graesser and Stanley P Franklin. 1990. Quest: A cognitive model of question answering. *Discourse processes*, 13(3):279–303.

Jian Guan, Xiaoxi Mao, Changjie Fan, Zitao Liu, Wenbiao Ding, and Minlie Huang. 2021. Long text generation by modeling sentence-level and discourse-level coherence. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 6379–6393.

Xinyu Hua and Lu Wang. 2020. PAIR: Planning and iterative refinement in pre-trained transformers for long text generation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 781–793, Online. Association for Computational Linguistics.

Mead Hunter. 2016. From conflict to concord: Lessons from the mouse. *Etudes*.

Daphne Ippolito, Reno Kriz, Maria Kustikova, João Sedoc, and Chris Callison-Burch. 2019. Comparison of diverse decoding methods from conditional language models. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3752–3762.

Arnav Jhala and R. Michael Young. 2010. Cinematic visual discourse: Representation, generation, and evaluation. *IEEE Transactions on Computational Intelligence and AI in Games*, 2(2):69–81.

Max Kreminski, Melanie Dickinson, Michael Mateas, and Noah Wardrip-Fruin. 2020a. Why Are We Like This?: Exploring writing mechanics for an AI-augmented storytelling game. In *Proceedings of the 2020 Conference of the Electronic Literature Organization*.

Max Kreminski, Melanie Dickinson, Michael Mateas, and Noah Wardrip-Fruin. 2020b. Why Are We Like This?: The AI architecture of a co-creative storytelling game. In *International Conference on the Foundations of Digital Games*.

Max Kreminski, Melanie Dickinson, and Noah Wardrip-Fruin. 2019a. Felt: a simple story sifter. In *International Conference on Interactive Digital Storytelling*, pages 267–281. Springer.

Max Kreminski and Michael Mateas. 2021. Reflective creators. In *International Conference on Computational Creativity*.

Max Kreminski, Ben Samuel, Edward Melcer, and Noah Wardrip-Fruin. 2019b. Evaluating AI-based games through retellings. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 15, pages 45–51.

Max Kreminski and Noah Wardrip-Fruin. 2019. Generative games as storytelling partners. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*.

Anne Lamott. 2007. *Bird by Bird: Some Instructions on Writing and Life*. Knopf Doubleday.

Antonios Liapis, Georgios N. Yannakakis, Constantine Alexopoulos, and Phil Lopes. 2016. Can computers foster human users' creativity? theory and praxis of mixed-initiative co-creativity. *Digital Culture & Education (DCE)*, 8(2):136–152.

Enrique Manjavacas, Folgert Karsdorp, Ben Burtenshaw, and Mike Kestemont. 2017. Synthetic literature: Writing science fiction in a co-creative process. In *Proceedings of the Workshop on Computational Creativity in Natural Language Generation (CC-NLG 2017)*, pages 29–37.

Chris Martens, Anne-Gwenn Bosser, Joao F Ferreira, and Marc Cavazza. 2013. Linear logic programming for narrative generation. In *International Conference on Logic Programming and Nonmonotonic Reasoning*, pages 427–432. Springer.

Chris Martens and Rogelio E. Cardona-Rivera. 2016. Generating abstract comics. In *Proceedings of the 10th International Conference on Interactive Digital Storytelling*, pages 168–175. Springer.

Chris Martens, Joao F Ferreira, Anne-Gwenn Bosser, and Marc Cavazza. 2014. Generative story worlds as linear logic programs. In *Seventh Intelligent Narrative Technologies Workshop*.

Lara Jean Martin. 2021. *Neurosymbolic Automated Story Generation*. Ph.D. thesis, Georgia Institute of Technology.

Michael Mateas and Andrew Stern. 2003. Façade: An experiment in building a fully-realized interactive drama. In *Game Developers Conference*, volume 2, pages 4–8.

Louise Ousby. 2009. *Whatever it takes: an exploration of writing tools and strategies for completing a novel*. Ph.D. thesis, Queensland University of Technology.

Radha O'Meara. 2015. Changing the way we think about character change in episodic television series. *Journal of Screenwriting*, 6(2):189–201.

Julie Porteous, Marc Cavazza, and Fred Charles. 2010. Applying planning to interactive storytelling: Narrative control using state constraints. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 1(2):1–21.

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

Elena Rishes, Stephanie M Lukin, David K Elson, and Marilyn A Walker. 2013. Generating different story tellings from semantic representations of narrative. In *International Conference on Interactive Digital Storytelling*, pages 192–204. Springer.

David L Roberts and Charles L Isbell. 2007. Desiderata for managers of interactive experiences: A survey of recent advances in drama management. In *Proceedings of the First Workshop on Agent-Based Systems for Human Learning and Entertainment (ABSHLE 07)*.

Melissa Roemmele and Andrew Gordon. 2018. Linguistic features of helpfulness in automated support for creative writing. In *Proceedings of the First Workshop on Storytelling*, pages 14–19.

James Ryan. 2018. *Curating Simulated Storyworlds*. Ph.D. thesis, University of California, Santa Cruz.

James Owen Ryan, Michael Mateas, and Noah Wardrip-Fruin. 2015. Open design challenges for interactive emergent narrative. In *International Conference on Interactive Digital Storytelling*, pages 14–26. Springer.

Ben Samuel, Michael Mateas, and Noah Wardrip-Fruin. 2016. The design of writing buddy: a mixed-initiative approach towards computational story collaboration. In *International Conference on Interactive Digital Storytelling*, pages 388–396. Springer.

Abigail See, Aneesh Pappu, Rohun Saxena, Akhila Yerukola, and Christopher D. Manning. 2019. Do massively pretrained language models make better storytellers? In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 843–861, Hong Kong, China. Association for Computational Linguistics.

Alireza Shirvani, Rachelyn Farrell, and Stephen G Ware. 2018. Combining intentionality and belief: Revisiting believable character plans. In *Fourteenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Alireza Shirvani, Stephen G. Ware, and Rachelyn Farrell. 2017. A possible worlds model of belief for state-space narrative planning. In *Proceedings of the 13th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*.

Ben Shneiderman. 2007. Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM*, 50(12):20–32.

Adam Marshall Smith. 2012. *Mechanizing Exploratory Game Design*. Ph.D. thesis, University of California, Santa Cruz.

Ingibergur Stefnisson and David Thue. 2018. Mimisbrunnur: AI-assisted authoring for interactive storytelling. In *Proceedings of the AAAI Conference on artificial Intelligence and Interactive Digital entertainment*, volume 14, pages 236–242.

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

Bowen Tan, Zichao Yang, Maruan AI-Shedivat, Eric P Xing, and Zhiting Hu. 2021. Progressive generation of long text with pretrained language models. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4313–4324.

Theo Wadsley and Malcolm Ryan. 2013. A belief-desire-intention model for narrative generation. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Yashen Wang, Huanhuan Zhang, Zhirun Liu, and Qiang Zhou. 2021. Hierarchical concept-driven language model. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 15(6):1–22.

Stephen G Ware and R Michael Young. 2011. Cpocl: A narrative planner supporting conflict. In *Seventh artificial intelligence and interactive digital entertainment conference*.

Stephen G Ware and R Michael Young. 2014. Glaive: a state-space narrative planner supporting intentionality and conflict. In *Tenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

R Michael Young. 1999. Notes on the use of plan structures in the creation of interactive plot. In *AAAI fall symposium on narrative intelligence*, pages 164–167.

R. Michael Young, Stephen Ware, Brad Cassell, and Justus Robertson. 2013. Plans and Planning in Narrative Generation: A Review of Plan-Based Approaches to the Generation of Story, Discourse, and Interactivity in Narratives. *Sprache und Datenverarbeitung*, 37(1–2):67–77.

Rolf A. Zwaan and Gabriel A. Radvansky. 1998. Situation models in language comprehension and memory. *Psychological bulletin*, 123(2):162.