The Authoring Tool Evaluation Problem

Charlie Hargood and Daniel Green

Bournemouth University, UK
{chargood,dgreen}@bournemouth.ac.uk

1 The Authoring Problem and Evaluation

If part of "The Authoring Problem" is our collective lack of understanding of interactive narrative authorship, then authoring technology may be one of the most poorly understood areas of all. For the purpose of this proposal, we define authoring tools as applications designed for use by writers or narrative designers principally for the creation of specifically interactive narrative. Thus this definition includes *Twine*, *StorySpace*, *Inform*, *articy:draft* and similar systems, but excludes IDEs (principally used by programmers rather than writers or narrative designers) and excludes word processors (potentially used by writers but not overall principally for the creation of interactive narrative).

A significant amount of interactive narrative is created using authoring tools, from major game studios such as *THQ* and *CD Projekt Red* using *articy:draft*, other studios such as *Larian* and (the now defunct, but notable for interactive narrative) *Telltale* using bespoke in-house tools, to classic hypertext fiction authors who made and make use of systems such as *HyperCard*, *StorySpace* [1], and *Twine*, and to the burgeoning interactive fiction and text adventure communities use of systems such as *Inform* and *Quest*. Authoring tools form a substantial part of research in this community with a wide diaspora of tools appearing from research communities such as *ICIDS* and *ACM Hypertext* like *StoryPlaces* [4], *ASAPS* [5], and the recent *Villanelle* [6], and workshop series such as *AIS* are dedicated to their study. Authoring tools are both a fundamental part of the creation of interactive narrative and a significant part of our field.

So, what makes for an effective authoring tool? How does authoring tool design impact the author experience? How does tool design impact the stories created in these tools? The answer to all of these questions is, largely, "we don't know". Despite authoring technology being a critical facet of interactive narrative, very few formal evaluations of the user experience and efficacy of authoring tools have been conducted, and those that have are often brief.

2 A Proposition for Exploration

We propose a chapter or chapters should explore this problem:

- **The Scale** — What tools have or haven't been evaluated and to what degree.
- **The Explanation** — What has caused this issue and what are the barriers.
- **The Solution** — Ways forwards to address the issue.

We hypothesize that a significant cause for this issue is the difficulty in evaluating authoring tools using conventional methods. The well-established UX evaluation best practice is often couched in web applications or other tools with short granular tasks for evaluation [2]. This has led to a form of task-based analysis over short user studies that is poorly suited for authoring tools. The creative processes in authoring interactive narrative can last many months if not years and involve a complex network of tasks and approaches that are interconnected. While we might conduct a 30 minute controlled user study for a simple web application or even playing part of a computer game, doing this for a creative process presents significant issues. Consequently, faced with the difficulty of collecting useful data on long creative processes with difficult to control for methods, we fear the community has been intimidated out of evaluating its tools and instead focuses on the evaluations of creative works, themselves which remains valuable and are arguably easier to overcome.

Greenburg [3] called for user study research to unshackle itself from established methods and methodological expectations to develop methodologies that suit the tools and context in question. He explained how a dogmatic approach to user evaluation both fails to capture the value of some tools, stifles innovation, and limits understanding of tools in new domains. We feel the solution to this aspect of the authoring problem is a revolution in authoring tool evaluation methodologies, where new bespoke methods are developed and trialed to better understand the authoring tools at the heart of our field. Such methods need to account for the scale and variety of the creative processes involved, but must also not be afraid to take risks in experimenting with new methods that are workable in order to break the deadlock in this part of our community. This means accepting compromise and constraints to evaluation, and a movement away from some traditional methods, while still prioritising rigor and trying to refine our methods and approaches as they mature.

References

1. Bernstein, M.: Storyspace 1. In: Proceedings of the thirteenth ACM conference on Hypertext and hypermedia. pp. 172–181 (2002)
2. Goodman, E., Kuniavsky, M., Moed, A.: Observing the user experience: A practitioner's guide to user research. Elsevier (2012)
3. Greenberg, S., Buxton, B.: Usability evaluation considered harmful (some of the time). In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 111–120. CHI '08, ACM (2008)
4. Hargood, C., Weal, M.J., Millard, D.E.: The Storyplaces Platform: Building a Web-Based Locative Hypertext System. In: Proceedings of the 29th ACM Conference on Hypertext and Social Media. HT '18, ACM, New York, NY, USA (2018)
5. Koenitz, H.: Extensible tools for practical experiments in idn: the advanced stories authoring and presentation system. In: ICIDS. pp. 79–84. Springer (2011)
6. Martens, C., Iqbal, O.: Villanelle: An authoring tool for autonomous characters in interactive fiction. In: ICIDS. pp. 290–303. Springer (2019)