Intelligent Narrative Technologies and Social Believability in Games:
Papers from the AIIDE 2015 Joint Workshop

Cinematic, Ambient, Inhabitable Narrative Environments: Story Systems in Search of an Artificial Intelligence Engine

Steven Wingate
Department of English
South Dakota State University
Brookings, SD 57007
steven.wingate@sdstate.edu

Abstract

Cinematic, Ambient, Inhabitable Narrative Environments (CAINEs) are conceptual AI-driven interactive story systems combining text, audio, and visual imagery that are scalable and adaptable to a wide range of storytelling needs and interactor inputs. Conceived by at artist outside the AI community, they represent an opportunity to use AI in a nontraditional and immersive narrative fashion that relies not on the goal-based arrangement of story elements, but on the accretion and association of those elements in the minds of interactors. This paper represents the initial phase of the project's development.

Concept and Theoretical Underpinnings

The atomization of narrative brought on by its increasing digitalization—through gaming, interactive fiction, database cinema, etc.—presents unique opportunities for storytellers who eschew traditional monolinear narrative practices and embrace procedural decision-making, computational chance, and multisensory interactor inputs. Cinematic, Ambient, Inhabitable Narrative Environments (CAINEs) are proposed scalable AI-driven responsive storytelling systems consisting of linguistic, visual, and audio stimuli, either authored specifically for a given system or harvested from existing data sets, that are presented as micro-narrative events. These stimuli are invoked using haptic, vocal, gestural, or biometric inputs to form fragmentary narrative environments, which immerse interactors in story experiences that one can call multiform or mulivariant (Murray 1997; Ryan 2004).

Much AI work in gaming and interactive digital narrative has gone into manipulating combinations of narrative events to form multiple plausible and coherent plots. This approach, which emphasizes the functionality and ordering of plot, is fundamentally Proppian, and indeed Vladimir Propp's morphological approach has influenced efforts in computational story management and generation from the 1970s to the last decade (Klein et al. 1976; Grabson and Braun 2001; Hartmann et al. 2005). CAINE systems demand instead a broad and flexible conception of what comprises the narrative experience, and they do not privilege plot functionality or the traditional expectations of coherence. They invoke instead what Marie-Laure Ryan calls 'storiness'—the varying degrees to which a given work may or may not have narrative qualities (2006). This concept leaves room for narrative experiences to take place within works that are not "narratives" *per se*.

As AI has been useful in shaping the development of Proppian story systems, it can also shape systems built on more "fuzzy" (to use Ryan's term) visions of narrative experience. CAINEs are built upon a geographical conception of narratives as liminal spaces to be explored, with the interactor fully invited into the process of meaning-creation. The idea of the interactor as co-creator goes back at least to John Dewey in the early 20th century (1934). Wolfgang Iser's reader/response theory presages digital storytelling in that it emphasizes the highly active role of the reader in constructing meaning (1978). AI engines can enable a kind of immersivity that is not dependent on plot, but relies on the meaning created by the interactor by means of associations made between narrative stimuli. (Because of the active investment required of the interactor, the word *inhabitable* replaces the broader term *immersive* here.) In this, the CAINEs owe their allegiance to a contemporary of Propp's, filmmaker and theorist Sergei Eisenstein, whose "collision" theory of montage is exceptionally applicable to interactive digital narrative (1923).

Embracing this cognitive construct leads to an increased freedom in the way story elements can be combined computationally, since meaning takes shape through the accretion, congruence, or collision of narrative stimuli rather than through their precise and plausible ordering. With the pressures of linking a traditionally coherent Aristotelian narrative removed (as well as the structure of achievement/reward systems), AI engines are freed to maximize the relationships between stimuli by managing their collisions and congruencies in fluid response to interactor input. Story systems built from this idea become what Coester calls "experience networks" (2015). In this way, experiencing a CAINE is akin to sinking into one's own memory.

Like memories, the stimuli contained within each CAINE are broad and fragmentary; narrative material is atomized into even smaller narrative units than Propp envisioned with his *motifs*. (The terms *stimuli* and *narrative environment* mark a distinction between Proppian, event-centered conceptions of story and those that flow from Ryan's idea of stori*ness*.) Narrative environments include not only the events of a story, but ambient stimuli—its setting, mood, history of characters and communities, etc.—as well as interactor inputs, which all engender unique story experiences for each user. The elements of such an environment are significantly more diverse than plot events, and expressible through a greater variety of media. Because they carry information that is not ostensibly plot-related, stimuli are free to engage interactors in asequential ways. Thus, narrative functionality comes to mimic the associational (and frankly chaotic) functionality of the human mind.

CAINEs are an attempt to embrace and utilize that chaos in the service of the narrative experience. Mittlbock, utilizing Winnicott's psychoanalytic concept of the "intermediate area of *experiencing*, to which inner reality and external life both contribute," posits that "[i]mmersion into an intermediate area always demands surrender to chaos" and that "[s]urrender is a precondition for immersion" (Winnicott 1971 and 2005; Mittlbock 2012) CAINEs function in precisely this liminal space, and their inhabitability is a function of an emotional identification with uncertainty.

System and Interface Dynamics

The CAINE concept is scalable and story systems can be developed for a broad range of interfaces, from solo "confessional" booths that respond to inputs from a single interactor at a time, to large-scale, museum-sized installations that respond to the density of a crowd. (Factors such as memory constraints, size of stimuli database, etc. will affect scale in each case.) Each system maximizes juxtaposition by asking interactors to navigate multiple, interlocked sets of stimuli, which may or may not result in a definitive endpoint. User interfaces would be specific to each CAINE, and may include vocal (e.g., voice recognition and text-to-speech tools), haptic (touchscreen) gestural (Leap Motion), and biometric (heart rate, density of occupants in a space) inputs. A CAINE harvests and processes interactor data, determining the sequence of stimuli either through direct, prompted commands (the spoken word "more" calls up related stimuli) or through analysis of free interactor input (a parser responds to the linguistic content of speech). This approach allows interactors to be aware of their role in shaping the narrative experience without feeling that they are in control of it—a sense of decontrol designed to increase inhabitability.

Conceptually, CAINEs are a direct descendant of the Electronic Visualization Laboratory's CAVE system (Cruz-Neira et al. 1992). They can be executed on a pre-existing system such as Mechdyne's Cave2 or in custom configurations. CAINE interfaces also build on developments in the field of Intelligent Personal Agents (IPAs), particularly Blast Theory's *Karen* app which is built for art and entertainment, in contrast to the productivity function of most IPAs (2015). Unlike productivity IPAs, CAINEs embrace the uncertainty of communication between interactors and the AI that pulls stimuli from its database. Each CAINE is developed from specific linguistic and audiovisual data sets consisting of still and moving images, speech, ambient sounds, and text that can be interactive or merely graphical. Relationships between these stimuli may not be consciously apparent to interactors, which makes uncertainty a narrative actor in itself.

Stimuli data can be authored specifically for a given CAINE or harvested from larger scale data sets, as several artists and researchers have explored (Gordon 2014; Li 2014; Swanson 2008**)**. The following examples are intended to give a sense of the concept's range and potential uses. Ideally, similar AI engines can drive all of them; what differs are the stimuli in their databases, the specifics of the relationships between those stimuli, and the kinds of interactor inputs harvested. Because no system has yet been developed to host them, no works following the CAINE model have been created to this point.

The Dessication of Joe. A Tuscon trader in stocks and drugs, Joe tries to escape the blowback of a bad deal by abandoning his family and secretly walking across the Sonoran desert, where he will meet a man who can sell him a new identity. But the man never shows, so Joe wanders the desert thinking of the family he left behind and the new self he might never achieve. Designed for a single interactor using custom-authored media assets, this CAINE is somewhat game-like in that it is possible to "save" Joe, who without interactor assistance will simply die in the desert. Stimuli include live desert point-of-view videos,

ambient desert audios, visual and audio flashbacks and flash-forwards, etc. Interactors navigate through the desert (and the stimuli) using gesture, and they give instructions to a parser in response to the AI engine's interrogations.

The Famous Author Museum. Designed for museum installation as an immersive introduction to the life and work of a famous author, this CAINE uses natural language processing to cull phrases from the author's *ouvre*, which are then either converted to audio via text-to-speech or shown graphically. Stimuli would also include video footage, photographs, archival manuscript materials, interviews with friends and colleagues, etc. Though set up to sequence through its stimuli algorithmically without interactor guidance, it can nonetheless respond to touch screen input by revealing more information on an indicated subject (e.g., touching a picture of the author's childhood home calls up more details about it). This use of the CAINE concept is essentially an AI-driven interactive documentary.

The Whispering Conspiracy. This multi-screen environment for galleries and museums harvests web and social media sentences that include the word "conspiracy" and displays them as moving graphic ephemera. These are paired with clips from romance and propaganda films in the public domain (some of them data-mashed) as well as custom-authored gossip and innuendo whispered among a small circle of friends and lovers. Less ostensibly interactive than most, this CAINE responds to interactor density in the exhibition space, which determines where stimuli will appear. In this way, interactors both lead and are led by the exhibit's AI.

Role and Nature of AI System

Given the amount and variety of stimuli stored in their databases, the CAINEs could easily fall into randomness. While it is human nature to try making narrative sense of such randomness, our sensibility best creates meaning with the aid of theme and structure. It is possible to hand-script connections between such diverse stimuli in a non-AI system, with video clips triggering audios that trigger text displays, etc. But such systems offer limited opportunity for meaning-making because their range of collisions are limited by static relationships between elements (e.g., the keyword relationships within the Korsakow cinema database system remain fixed regardless of how they are used). Compared to the goal of the CAINEs, their interactivity is shallow because they respond only in predetermined ways (i.e., stimuli begetting only specific other stimuli) that do not respond sufficiently to interactor input. AI systems could accommodate this need for responsiveness.

An AI architecture for the CAINEs would allow a far more robust set of relationships between stimuli, ranging from fixed connections to opportunities for randomization, that are determined in real time. Riedl's concept of the *experience manager* well describes the AI engine that will reside behind the CAINEs. Managers built for this purpose avoid two constraints common to AI practice and research. The first is the need for a coherent narrative, which Riedl describes as "one in which all events build off prior events until a conclusion is reached." (2013) Because CAINEs provide an inhabitable rather than a sequential narrative, this need is subverted. Secondly, there is what Magerko calls the *boundary problem*—"player actions bringing a dramatic experience outside of the boundaries of authored content" (2005). This is sidestepped because the CAINEs are designed so that all choices within a given system, and all collisions between stimuli, are possible. (The specific narrative elements within each system of course constrain choice and define the precise AI solutions required. CAINE systems do not need to be AI complete to function, and the particular experience management required will depend on technical and aesthetic factors.)

The CAINEs, then, offer an opportunity to see how AI experience managers can guide the human mind through a narrative environment when their computational mechanics are free to work through its possibilities without the need to facilitate the navigation of multiple coherent stories. The diverse and fragmented nature of the CAINE stimuli, as well as the way those stimuli are navigated, makes for a narrative experience that is largely assembled in the interactor's unconscious, and the AI system required for it will essentially mimic unconscious processes. The procedural nature of computation makes strange bedfellows of AI and the very idea of the unconscious. But disengaging the tools of digital storytelling from traditional narrative sense-making, as CAINEs do, allows AI engines to function as a "computational subconscious," which can offer insight into the process of meaning-making at the moment it begins to take shape in the mind. Through cognitive psychology models—perhaps especially in fields that explore unconscious cognition—computational expression of human unconscious processes may be possible. CAINEs are both a laboratory and playground for using AI to explore that expression.

Precedent Media Systems

CAINEs can potentially deliver any textual, audio, or visual stimuli; shaping a particular system involves creating unique sets of relationships between stimuli, AI engines, and potential responses to interactor input. The CAINE concept seeks to build on two media projects that share parts of its approach and offer encouraging precedents.

One clear precedent for the CAINEs (and a principal inspiration for them) is the database-driven ambient video

system developed by Jim Bizzocchi (2008). This system contains databases of video clips (without sound) and of transitions, and is intended for ambient viewing; it proceeds algorithmically whether or not viewers are engaged. Created using MaxMSP/Jitter video performance software and JavaScript, its proof-of-concept project involves nature imagery. The system can also accommodate other kinds of content, and an algorithmic "city symphony" is in planning. The system may eventually be developed for potential third-party use and be adaptable for CAINE interfaces.

Also directly applicable to the CAINEs, and even more significant in establishing their theoretical and practical path, is the concept and practice of enactive media (Kaipainen et al. 2011). This approach

> conceives the underlying technology as continuous, ubiquitous and "intelligent" accompaniment to the human actor, or a direct extension of the user's perceptual and cognitive apparatus involved in participation in the system…. Enactive media assume enactive systems as their core, but also a repertoire of content elements that can be used to generate a range of meanings in terms of narrative recombinations in real time. The elements can be pre-recorded film footage, audio tracks or text excerpts or, alternatively, real-time generated behaviors.

Enactive systems use psycho-physiological data collection tools and a database of stimuli to form a feedback loop between the viewer and video, using a computational "montage machine"—a function performed in the CAINEs by an AI engine. This work (exemplified by Pia Tikka's *Obsession*) points not only to ways that interactor input can be harvested and channeled, but to ways that cognitive and computational models can be utilized in the arrangement of stimuli. The CAINE project will closely track the further work of these researchers.

Project Development

Since this concept comes *to* the AI community, rather than *from* it, making the CAINEs a reality will require AI expertise beyond this author's scope. The project seeks guiding partners who can:

- Explore AI architecture precedents, with the goal of devising systems to accommodate the project's data management needs;
- Identify cognitive and computational models that can be utilized in the creation of interactor-responsive algorithms;
- Develop flexible AI engines that can power CAINEs of various sizes and scopes;
- Establish appropriate scale and expectations for a proof-of-concept CAINE prototype, with the proximate goal of museum or gallery installation;
- Recruit other authors to work in the system once prototyping is complete.
- Document and actively publish the project's progress as an alternative use of AI in interactive digital narrative.

References

Bizzocchi, J. 2008. Ambient Video: The Transformation of the Domestic Cinematic Experience. In Hawk, B.; Rieder, D.; and Oviedo, O. (eds.) *Small Tech: The Culture of Digital Tools*. Minneapolis: University of Minnesota Press.

Blast Theory 2015. Karen: An App. Artwork. http://www.blasttheory.co.uk/projects/karen/.

Coester, D. 2014. The Reverbatory Narrative: Toward Story as a Multisensory Network. In Grabowski, Michael (ed.) *Neuroscience and Media: New Understandings and Representations*. New York: Routledge.

Cruz-Neira, C.; Sandin, D.; DeFanti, T.; Kenyon, R.; Hart, J. 1992. The CAVE: Audio Visual Experience Automatic Virtual Environment. *Communications of the Association for Computing Machinery* 35(6): 64-72.

Dewey, J. 1934, 2005. *Art as Experience*. New York: Putnam.

Eisenstein, S. 1923. Montage of Attractions. *LEF* (Moscow). Tr. by Daniel Gerould. 1974. *The Drama Review* 18(1): 77-85.

Grasbon, D. and Braun, N. 2001. "A Morphological Approach to Interactive Storytelling." In Proceedings of Artificial Intelligence and Interactive Entertainment, CAST '01, Living in Mixed Realities, Sankt Augustin, Germany.

Gordon, A.; Wang, C.; Bosack, M.; Sininger, J. 2014. Civilian Analogs of Army Tasks. Artwork. Exhibition of 7th International Conference on Interactive Digital Storytelling, Singapore.

Hartmann, K., Hartmann, S. and Feustel, M. 2005. Motif Definition and Classification to Structure Non-linear Plots and to Control the Narrative Flow in Interactive Dramas. In Subsol, G. (ed.) *Virtual Storytelling Using Virtual Reality Technologies for Storytelling*, Third International Conference, ICVS 2005, Lecture Notes in Computer Science 3805 Springer, 158-167.

Iser, Wolfgang. *The Act of Reading: A Theory of Aesthetic Response*. Johns Hopkins University Press, Baltimore and London, 1978.

Kaipainen, M; Ravaja, N.; Tikka, P.; Vuori, R.; Pugliese, R.; Rapino, M; and Takala, T. 2011. Enactive Systems and Enactive Media: Embodied Human-Machine Coupling beyond Interfaces. *Leonardo* 44(5) 433–438.

Klein, S., Aeschlimann, J. F., Applebaum, M.A., Balsiger, D.F., Curtis, E.J., Foster, M., Kalish, S.D., Kamin, S.J., Lee, Y.-D. and Price L.A. 1976. Simulation d'hypothèses émises par propp et Lévi-strauss en utilisant un système de simulation meta-symbolique. Informatique et Sciences Humaines (28): 63-133.

Li, B.; Thakkar, M.; Wang, Y.; Riedl, M. 2014. Storytelling with Adjustable Narrator Styles and Sentiments. In Proceedings of 7th International Conference on Interactive Digital Storytelling. Lecture Notes in Computer Science, Volume 8832. Berlin: Springer.

Magerko, B. 2005. Evaluating Preemptive Story Direction in the Interactive Drama Architecture. *Journal of Game Development* 2(3): 25–52.

Mittlbock, K.S. 2012. Personality Development Through Immersion into Intermediate Areas of Digital Role-Playing Games. In

Computer Games and New Media Cultures. A Handbook of Digital Games Studies. Amsterdam: Springer.

Murray, J. 1997. *Hamlet on the Holodeck*. Cambridge: The MIT Press.

Riedl, M. and Bulitko, V. 2013. Interactive Narrative: An Intelligent Systems Approach. *AI Magazine* 34(1): 67-77.

Ryan, M-L. 2004 Multivariant Narratives. In Schreibman, S.; Siemens, R.; Unsworth, J. (eds.) *A Companion to Digital Humanities*. Oxford: Blackwell.

Ryan, M-L. 2006. *Avatars of Story*. Minneapolis: University of Minnesota Press.

Swanson, R; Gordon, A. 2008. Say Anything: A Massively Collaborative Open Domain Story Writing Companion. In First International Conference on Interactive Digital Storytelling. Lecture Notes in Computer Science, Volume 5334. Berlin: Springer.

Tikka, P. 2005. *Obsession*. Artwork. Oblomovies Film Productions, Helsinki.

Winnicott, D.W. 1971, 2005. *Playing and Reality*. New York: Routledge.