A Discourse-first Approach to Visual Narrative Generation through Operationalizing Comic Panel Transitions
Paper type: System Description

Chris Martens
Expressive Intelligence Studio
Computational Media Department
University of California, Santa Cruz
Santa Cruz, CA, USA
crmarten@ucsc.edu

Rogelio E. Cardona-Rivera
Liquid Narrative Group
Department of Computer Science
North Carolina State University
Raleigh, NC, USA
recardon@ncsu.edu

Abstract

This abstract is so awesome.

Introduction

- What are we trying to do?
- What is our approach?
- Talk about how creative the discipline is
- Why does discourse-driven help creativity in narrative generation?
  - Creative potential is bigger in discourse-driven approach, principle of least-commitment in the fabula - don't need to simulate a portion of the narrative universe which is never revealed which may impose limitations on future narrative generation. **[RCR:** *Example!***]**
  - Focusing on the telling may leave aspects of fabula unspecified, which may broaden the interpretation of the story in the minds of story consumers. **[RCR:** *Example!***]**
  - Talk about the pipeline model of narrative generation (primarily simulation focused)
  - We're exploring an alternative account - focus on the telling of the story, let story consumers "fill in the gaps"
    - Gricean Maxims
    - Closure principle Saraceni - Third aspect of "relatedness", depends on not what is overtly explicit in the narrative's surface code, but also on inference. "Closure" comes from Gestalt principles.
    - Gestalt principle
- Why are comics a great domain for computational creativity?

We pursued a discourse-driven

System Description

Our approach to generating visual narratives begins as a linear process that selects next comic panels based on the contents of previous panels, choosing randomly among indistinguishably-valid choices. The concepts we represent formally are *transitions*, *frames*, and *visual elements*, which we define below.

A **visual element (VE)** is a unique identifier from an infinite set, each of which is possible to map to a distinct visual representation. We do not explicitly tag visual elements as specifically characters, props, or scenery, making the representation agnostic to which of these narrative interpretations will apply. In the visual rendering of our comics, we represent VEs as random combinations of shape, color, and size, supplying additional inputs to the human cognitive processes that may interpret these elements' narrative role.

A **frame** is a panel template; at the abstract generation level, it includes an identifier or set of tags and a minimum number of required visual elements. The reason a frame specifies a *minimum* number of VEs is to allow for augmentation of the frame with pre-existing elements: for example, the *monologue* frame requires at least one visual element, indicating a single, central focal point, but other visual elements may be included as bystanding characters or scenery elements. At the rendering level, a frame includes instructions for where in the panel to place supplied visual elements.

A **panel** is a frame instantiated by specific visual elements.

Finally, a **transition** is (XXX continue)

Formal Transition Types

Moment: keep VEs, change frame and/or modifiers.

Add: introduce a VE that didn't appear in the previous panel (but might have appeared earlier).

Subtract: remove a VE from the previous panel (and potentially choose a new frame).

Meanwhile: choose a new frame and only show VEs that didn't appear in the previous panel, generating new VEs if necessary.

Rendez-vous: choose a new frame and fill it with any combination of previously-appearing VEs. Generate new VEs only when there aren't enough previous VEs to fill the frame.

Example

Implementation

Output

Related Work

Historically, the computational generation of narrative has followed what Ronfard and Szilas (2014) term the *pipeline model*: a narrative artifact is computationally generated by first simulating the story world as a collection of events, and then piping the story world information to a discourse generator, which generates a selective presentation of story world events in a particular medium. Current work in the computational creativity community has primarily pursued this pipeline model for narrative generation. Work by...
[**RCR:** *There are several papers to cite here, but the gist is: pipeline model is pervasive.*]

As Ronfard and Szilas argue, the pipeline model is neither *necessary* nor *sufficient* for the successful generation of narrative structure. This is because authors intentionally design their narratives to affect audiences in specific ways (Chatman 1980; Bordwell 1989), which involves reasoning beyond what is communicated (the underlying story world) but rather how it is communicated. It is unnecessary to simulate an aspect of the narrative universe that is never communicated to the audience, if it does not inform the ultimate delivery of the narrative artifact. Similarly, it is insufficient to reason about the story and discourse constituents independent of each other, because the characteristics of a discourse realization shape the stories that can be told in that medium (Herman 2004). As mentioned earlier, constraining the generation process to the pipeline model unnecessarily restricts how creative the generator can ultimately be, since story world commitments are not revisited when generating discourse.

Our work presents a departure from the pipeline model, opting instead for a *discourse-first approach to narrative generation*. In this model, the story world is simulated inasmuch as is necessary to support the telling of story events in the discourse. The work we present here is a first step in this discourse-driven model, focused on understanding how the discourse of visual language narratives enforces constraints on the underlying story worlds they represent, and how these can further guide subsequent choices for discursive presentation.

- Talk about Understanding Comics (McCloud 1993)
- Talk about Visual Language of Comics (Cohn 2013)
- Talk about the MEXICA System (Pérez y Pérez and Sharples 2001) and how we're different
- Talk about the departure from traditional narrative generation work
  - Talk about the pipeline model of narrative generation (primarily simulation focused)
  - We're exploring an alternative account - focus on the telling of the story, let story consumers "fill in the gaps"
    - Gricean Maxims
    - Closure principle Saraceni - Third aspect of "relatedness", depends on not what is overtly explicit in the narrative's surface code, but also on inference. "Closure" comes from Gestalt principles.
    - Gestalt principle

Acknowledgments

These acknowledgments are tubular.

References

Bordwell, D. 1989. *Making Meaning: Inference and Rhetoric in the Interpretation of Cinema*. Cambridge: Harvard University Press.

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

Cohn, N. 2013. *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London, England, UK: Bloomsbury.

Herman, D. 2004. Toward a transmedial narratology. In Ryan, M.-L.; Ruppert, J.; and Bernet, J. W., eds., *Narrative Across Media: The Languages of Storytelling*. University of Nebraska Press. 47–75.

McCloud, S. 1993. *Understanding Comics: The Invisible Art*. New York, NY, USA: Harper Collins.

Pérez y Pérez, R., and Sharples, M. 2001. MEXICA: A computer model of a cognitive account of creative writing. *Journal of Experimental and Theoretical Artificial Intelligence* 13(2):119–139.

Ronfard, R., and Szilas, N. 2014. Where story and media meet: computer generation of narrative discourse. In *Proceedings of the 5th Workshop on Computational Models of Narrative*, 164–176.