Computational approaches to analyzing and generating comics

Aditya Upadhyayula (supadhy6@jhu.edu)
Department of Psychological and Brain Sciences, Johns Hopkins University, USA

Chris Martens (martens@csc.ncsu.edu)
Department of Computer Science, North Carolina State University, USA

Rogelio E. Cardona-Rivera (rogelio@cs.utah.edu)
School of Computing and the Entertainment Arts and Engineering Program, University of Utah, USA

Andrew T. Hendrickson (A.Hendrickson@uvt.nl)
Department of Cognitive Science and Artificial Intelligence, Tilburg University, The Netherlands

Neil Cohn (neilcohn@visuallanguagelab.com)
Department of Cognition and Communication, Tilburg University, The Netherlands

Keywords: comics; corpus linguistics; computational linguistics; computational modeling; narrative generation

Introduction

The past decade has seen an increasing focus on visual narratives like comics as an area for investigating numerous facets of cognition across numerous subfields of Cognitive Science (Cohn & Magliano, 2019). While early work focused on applying linguistic theories to analyze the structure of these visual and multimodal narratives, empirical work has extended to methods in cognitive psychology and cognitive neuroscience. This research has illustrated how the visual representations of sequential images share structural properties with language, and often overlap in their neurocognitive comprehension mechanisms (Cohn, 2019), albeit manifested in the visual-graphic modality rather than the verbal modality.

Adjacent to this has been a growing focus on computational methods applied to comics (Augereau, Iwata, & Kise, 2018; Laubrock & Dunst, 2019). These include the use of computational modeling to analyze corpora of comics, the use of parsers to extract underlying properties of comics and their comprehension, and the programming of computational systems to generate novel comics. These approaches again provide opportunities to integrate various facets of cognition, as they combine analytical methods often used to analyze or generate text alongside those often applied to visual representations.

This symposium surveys a range of these computational methods applied to the visual and multimodal properties of comics.

A corpus analysis of framing structure of panels and their cohesive links

A first way that computational analyses can inform research on visual narratives is through analysis of annotated comics (Laubrock & Dunst, 2019). There is at least tacit acknowledgement that visual narratives differ across cultures, though they may share basic underlying features. Yet, corpus analyses of the properties of such cross-cultural systems remains limited. In this presentation, ***Andrew Hendrickson*** will discuss an empirical analysis of annotated comics from the Visual Language Research Corpus (VLRC: http://www.visuallanguagelab.com/vlrc), which includes annotation of the properties of 290 comics from the United States, Europe, and Asia and contains over 36,000 comic panels. This analysis focuses on using sequential pattern mining (Zaki, 2001) to identify patterns in the framing of characters within panels and dimensions of meaningful relations between images (shifts in time, characters, or spatial location) both within and between panels. Subsequently, the similarity between comics across these sequential patterns are compared to both geographical and temporal distance using representational similarity analysis techniques (Kriegeskorte, Mur, & Bandettini, 2008). The results indicate that storytelling patterns in comics across the world differ in distinctly characterizable systems of visual narratives that emerge and differ over time.

Hierarchical structure in processing visual narratives: A computational investigation

Recent years have seen an increase in computational analyses of linguistic corpora and psycholinguistic data to explore the underlying structure and compositionality of language. In this presentation, ***Aditya Upadhyayula*** extends these methods to analyze a self-paced reading experiment of visual narrative sequencing. Participants viewed 6-panel long wordless comic strips panel-by-panel at their own pace where each sequence had two primary constituents. In experimental sequences, blank white "disruption" panels were either inserted within or between the constituents (Cohn, Holcomb, Jackendoff, & Kuperberg, 2012), where participants were found to spend longer time looking at the disruption panels when they were presented Within Constituents (WC) than when presented in Between Constituents (BC) of the

narrative structure. We then validated these results using two computational models. We compared the predictions of a probabilistic Earley parser (Hale, 2001) and a Hidden Markov Model (HMM) against the behavioral RT measures. While both the models output the surprisals for each panel in the comic strip, the Earley parser used information about the underlying narrative structure and the HMM used the category level bigrams in the comic strips to generate these surprisals. If our perception of comic strips is structured, then we should expect a negative correlation between the parser's outputs and the disruption panels' RTs when presented within (WC) and between (BC) constituents. We found a negative correlation as hypothesized. More specifically, the probabilistic Earley parser's outputs correlated more with the disruption panel RTs than the HMM model. Taken together, these results are consistent with findings that visual narratives use hierarchical structure, and they demonstrate that such methods from computational linguistics can provide insights into the formal structure of domains beyond verbal language.

Generation of visual narrative

Computational methods have also provided insight beyond analyses of existing data, but instead for their capacity to model the generation of expressions, including narrative. In interactive media, the problem of narrative generation supposes a computationally modeled universe in which the goal is to algorithmically generate narratively plausible trajectories of character behavior that meet authorial constraints. The last several decades have seen innovations in a "pipeline" approach to narrative generation, in which AI planning is used to construct plot structures and then discourse, or the telling of the story, is developed as a second pass. However, this approach poses challenges for the highly visual medium of comics, wherein the affordances for discourse shape the possibility space of available plots. Cognitive science indicates, for instance, that the layout of panels on a page, choices of framings within a panel, and careful elision of information are critical to how comic readers form mental models of the underlying story-world and plot of a comic. In this talk, ***Chris Martens*** and ***Rogelio Cardona-Rivera*** will discuss ongoing research on an alternative approach to computational narrative generation in which the cognition of comic readers directly drives a process for generating visual narratives.

Automatically generating news comics

Finally, generating comics can aid in applied settings. Because of their communicative power, comics can provide an effective method of communication for conveying news for less engaged members of the public, and indeed many forays into non-fiction comics have been made in the past decades. ***Neil Cohn*** describes a collaboration with BBC News Labs to create a tool that will help create "news comics" for wider accessibility for journalism. Starting with stories related to health, this system allows journalists to type a simple news story, and then organizes this text into a structured story template. It then pairs this text with images which are stored in an established visual vocabulary of an "image library." Images are 'tagged' with keywords describing that content. When those keywords appear in the inserted text, the system automatically links appropriate images to that text within a comic panel's frame. Simple content might involve concrete and specific elements like prominent societal figures (via names), places and locations (via names), organizations and companies (perhaps depicted with logos), and generic stereotyped professions (i.e., a doctor, a scientist, etc.). More complex image generation involves metaphors, whereby images are encoded with keywords associated to a metaphoric frame. For example, images of rainclouds might be encoded with keywords associated with their literal depictions (clouds, rain, storm) and with words associated with a metaphoric frame (such as "depression is a storm"). Altogether, this system allows for considerable generativity of images in combination with text to create an effective tool for multimodal expressions in communicating news.

References

Augereau, O., Iwata, M., & Kise, K. (2018). A survey of comics research in computer science. *Journal of Imaging, 4*(87), 1-19.

Cohn, N. (2019). Your brain on comics: A cognitive model of visual narrative comprehension. *Topics in Cognitive Science, 12*(1). doi:10.1111/tops.12421

Cohn, N., Holcomb, P., Jackendoff, R., & Kuperberg, G. (2012). *Segmenting visual narratives: evidence for constituent structure in comics.* Paper presented at the 34th Annual Conference of the Cognitive Science Society, Sapporo, Japan.

Cohn, N., & Magliano, J. P. (2019). Editors' Introduction and Review: Visual Narrative Research: An Emerging Field in Cognitive Science. *Topics in Cognitive Science, 12*(1). doi:10.1111/tops.12473

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, *2*, 4. doi:10.3389/neuro.06.004.2008

Laubrock, J., & Dunst, A. (2019). Computational approaches to comics. *Topics in Cognitive Science, 12*.

Zaki, M. J. (2001). SPADE: An efficient algorithm for mining frequent sequences. *Machine learning*, *42*(1-2), 31-60.

Hale, J. (2001, June). A probabilistic Earley parser as a psycholinguistic model. In Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies (pp. 1-8). Association for Computational Linguistics.