

Chapter 8

Can Computers Understand Picture Books and Comics?

Miki Ueno

Toyohashi University of Technology, Japan

Kiyohito Fukuda

Osaka Prefecture University, Japan

Naoki Mori

Osaka Prefecture University, Japan

ABSTRACT

The objective of this chapter is to develop a method for generating and analyzing creative work by using computers. In this chapter, picture books and comics are considered representative creative work because they contain multi-modal information represented by natural language and pictures. These works focus on interesting issues, which can be explained by computational approaches to narratology. In this chapter, the authors discuss the following two topics. First, a method of semi-automatic picture book generation by agent-based simulation is presented. Second, methods of generating and analyzing comics on the basis of the features of pictures and stories used are described. From that work, the authors introduce the construction of an original dataset.

INTRODUCTION

How do humans produce creative work? What do individuals imagine, and what do they think of during the creative process? The ultimate purpose of this study is to find an algorithm for the generation of creative, intellectual works, such as novels, comics, and animations, with regard to artificial intelligence. The creative works listed above involve stories and are constructed using sequential components. Further, the popular representation of a narrative utilizes natural language and pictures. For example, novels are constructed using natural language, and this type of language is a very useful means of logically depicting a given situation. However, the number of sentences comprising individual scenes varies considerably, and it is therefore difficult for computers to determine scene boundaries. In addition, comics are con-

DOI: 10.4018/978-1-5225-7979-3.ch008

Can Computers Understand Picture Books and Comics?

structured using both pictures and natural language, and animations are comprised of pictures and sound based on natural language. Note that pictures constitute an effective means of creating stories, because they can convey postures and positions more directly than verbal or written language. This approach allows stories to become comprehensible to a wider audience, regardless of age or nationality.

In Japan, comics and animations are common in popular culture. A key characteristic of these representations are the deformation or exaggeration of characters and other objects in the story. These visual representations contain lyrical and descriptive aspects to depict each situation clearly. Further, these works are often very intellectual, and some popular authors have produced exceptional creative pieces.

Story creation is something of which the vast majority of humans are capable; however, computers can neither create nor understand stories. The authors are very interested in the process of story construction, particularly as regards the development of appropriate models to facilitate computer story creation. Therefore, our goal is to define complete models of story creation for use by computers. However, it is difficult to define models for the entire process and every objective, therefore the authors suggest models for a specific process of creation. Hence, the authors have conducted several studies regarding computational creative stories, focusing on the topics listed below.

- **Story Generation and Picture Book Generation:** Semi-automatic story generation using log data.
 - Continuous transitions are automatically created and are given to writers in order to write unexpected stories that cannot be created by humans only. The aim of this section is to propose a computational method of coherent and unexpected story creation. and various picture book generation.
- **Comic Generation:** Automatic comic generation focusing on the relationship between transitions of stories and expression of pictures and comic analysis based on two of their features.
 - Continuous transitions expressed by pictures are modeled and the relations between characteristic expressions and the patterns of stories are analyzed. The aim of this section is to suggest computational models for comics focusing on the expressions in pictures.

In this chapter, these studies are introduced using examples, in order to demonstrate the fundamental techniques of computational narratology and the challenges involved in the quantification of story creation.

This chapter is enhanced one of “Can Computers Create Comics and Animations?” (Ueno, Fukuda, & Mori, 2016). Compared to our previous chapter, the authors introduce three major enhanced topics in this chapter.

- Picture Book Generation
- Analyzing Four-Scene Comics by Deep Learning Methods
- Constructing Original Four-Scene Comics for Machine Learning

BACKGROUND

Recently, automatic story generation and comprehension has attracted considerable interest as a challenging problem in the fields of artificial intelligence. The types of stories are broadly classified into two categories. One is utilizing natural language, the other is utilizing pictures. Novels are written in

31 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/can-computers-understand-picture-books-and-comics/221858

Related Content

Kabuki as Multiple Narrative Structures and Narrative Generation

Takashi Ogata (2019). *Post-Narratology Through Computational and Cognitive Approaches* (pp. 192-275).

www.irma-international.org/chapter/kabuki-as-multiple-narrative-structures-and-narrative-generation/221854

Understanding Health Communication in the Era of Media Convergence

B. K. Ravi (2024). *Changing Global Media Landscapes: Convergence, Fragmentation, and Polarization* (pp. 149-158).

www.irma-international.org/chapter/understanding-health-communication-in-the-era-of-media-convergence/350834

Internationally Educated Mathematics Teacher Candidates: Speaking About Mathematics, Language, and Society

Dragana Martinovic (2017). *International Journal of Semiotics and Visual Rhetoric* (pp. 36-46).

www.irma-international.org/article/internationally-educated-mathematics-teacher-candidates/183639

Social Media-Based Visual Strategies in Tourism Marketing

Jing Geand Ulrike Gretzel (2018). *International Journal of Semiotics and Visual Rhetoric* (pp. 23-40).

www.irma-international.org/article/social-media-based-visual-strategies-in-tourism-marketing/221148

Health-Related Communication and Rare Diseases: A Passport for the Patient Journey

Michaela Liuccio (2020). *Communicating Rare Diseases and Disorders in the Digital Age* (pp. 87-112).

www.irma-international.org/chapter/health-related-communication-and-rare-diseases/248321