Evaluation of a Narrative Discourse Generation System Based on the Concept of "Norm and Deviation"

Taisuke Akimoto
Graduate School of Informatics and Engineering, the University of Electro-Communications, 1-5-1 Chofugaoka Chofu, Tokyo 182-8585, Japan

Takashi Ogata
Faculty of Software and Information Science, Iwate Prefectural University, 152-52 Sugo Takizawa, Iwate 020-0693, Japan
E-mail: t8akimo@yahoo.co.jp, t-ogata@iwate-pu.ac.jp

Abstract

This paper deals with the verification of the narrative discourse system that automatically produces a variety of "discourse" structures from an inputted "story" structure through an iterative mutual action between a "narrator" mechanism and a "narratee" mechanism. In particular, we analyze a series of 10000 generated discourse structures according to their structural feature values by focusing on the diachronic alternation of "norm", the narratee's expectation in receiving discourses. Based on the results, we discuss the achievement and issues to be addressed.

Keywords: narrative generation system, automatic analyses, narrative discourse, norm and deviation, Jauss.

1. Introduction

Automatic narrative generation is a challenging topic in the field of artificial intelligence. In this topic, methodology for evaluating systems is a difficult issue and has addressed by several researchers. Rowe et al.[1] and Zhu[2] argued for the need to combine the multiple aspects including the authorial process, generated texts, and reading process in the evaluation of narrative generation systems. Pérez y Pérez[3] and Peinado et al.[4] tried to formalize such evaluation criteria as coherence, interestingness, and novelty in generated narratives.

We have been addressing the development of the "Integrated Narrative Generation System" (INGS) based on an "expanded literary theory", an interdisciplinary approach to narrative generation mechanism in INGS across informatics and literary theories.[5, 6] This paper deals with the verification of the narrative discourse system that we have developed as a practice of the expanded literary theory.[7, 8]

The system automatically produces a series of "discourse" structures from an inputted "story" structure through an iterative mutual action between a "narrator" mechanism, which generates discourse structures, and a "narratee" mechanism, which receives the generated discourses. This cyclic generation model continuously produces discourse structures through the diachronic alternations of "norm", the narratee's "expectation" in receiving discourse structures, via its "deviation" by the narrator. The term "norm" means a fixed frame at the time of generating discourse structures and "deviation" means an action to try to produce a new type of discourse structure by breaking the norm.

This paper objectively consider the behavior of the implemented system for identifying its achievements, limitations, and issues to be addressed. We use

quantitative analyzing methods based on a conceptual framework of "norm and deviation".

2. An Outline of the Narrative Discourse System

The narrative discourse system was implemented with Common Lisp. It automatically produces a series of discourse structures from an input story structure. The details of the mechanism and the generation examples were shown in our previous papers.[7, 8]

2.1. *The discourse structure generation*

The generation of a discourse structure in the system means a transformation from a story structure into a discourse structure. A story is a content of a narrative or a temporal sequence of events. In contrast, a discourse means a structure of how the story is narrated. Each of a story and a discourse is represented with a tree structure. The main element of the tree structure is "event", a character's action, at each terminal node. Each internal node is a "relation" among the child nodes.

The structural transformation is done by using "discourse techniques" which define transformational operations of a part of the structure. The techniques are defined by referring a part of the discourse categories by Genette.[9] 13 types of discourse techniques including temporal ordering, repetition, and so on are implemented in the system. Different discourse structures are generated according to what techniques are used and where the techniques are applied in the tree structure. The techniques to be used are determined based on the narrator's "generative goal" which we will describe latter. On the other hand, since the target of each technique is decided at random with several conditions, the output structures by a same generative goal have relatively small differences.

2.2. *The generation cycle*

The output discourse structures gradually change through the mutual action between the narrator and the narratee, called "generation cycle". In each step, the narrator generates a discourse structure from a story structure according to a set of parameters as "generative goal" or targeting structural features. On the other hand, the narratee feeds back an evaluation of the generated discourse according to a set of parameters as "expectation" or desiring structural features. The ten parameters corresponding to structural features relevant to the 13 discourse techniques are commonly used in the generative goal and the expectation: supplement, complexity, suspense, length, hiding, descriptiveness, repetition, diffuseness, implication, and temporal-independency. Each parameter takes a value of 1 (small), 2 (medium), or 3 (large).

The diachronic change of output discourses is arisen from the change of parameters' values in both the generative goal and expectation. This mechanism is modeled based on our original reinterpretation of a part of the literary history model by Jauss[10]. The narrator basically sets the generative parameters to fit the narratee's expectation and generates discourse structures iteratively. The narratee increases his satisfaction by receiving the fitted discourses to the expectation. The process, however, eventually reaches a point where the narratee gets tired or his satisfaction begins to fall. The turning point of the satisfaction is arbitrary set by the variable n_p. When this happens, "deviation" occurs and the narrator abandons a portion of the old generative parameters and moves to a new cycle of discourse grounded on the newly found strategy. The narratee's expectations change according to the reconstruction. In this generation model, the role of the narratee's expectation is to hold a "norm" for discourse generation.

3. The Framework of the Analyses

This section describes the framework of the analyses for a series of discourses produced by the system.

3.1. *The structure of a series of discourses*

A series of discourses can be structured as Fig. 1. Each allow beside "generative goal" and "expectation" in the figure means the duration of same parameters' values. A "shift in norm" means a change of a parameter's value in the expectation. We use a subscript number to indicate a specific norm like "$norm_i$". We call a discourse generated by the generative goal equal to the expectation "normative discourse" and a discourse by a generative goal not equal to the expectation "deviated discourse". The processing of "deviation" is to change a parameter's value in the generative goal at random and it causes a "shift in norm" in the next step. By segmenting the series with each "shift in norm", the tale discourse in each segment is the "deviated discourse" and the others are "normative discourses".

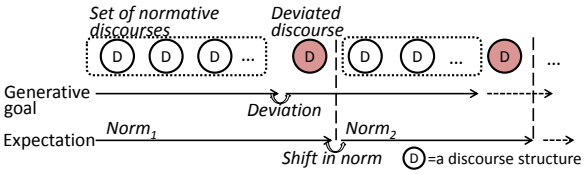

Fig. 1. The structure of a series of discourses.

3.2. *Four aspects of the analyses*

The basic idea of the analyses is to treat each discourse structure as a set of numerical values which represent structural features corresponding to the ten parameters. We call the values "discourse feature values" (DFVs). Each DFV is automatically calculated from a generated discourse. For example, "length" and "complexity" are respectively calculated based on the number of terminal nodes in a discourse structure and the number of relations needed for defining temporal order transformation. A set of discourse structures can be spacially treated based on their DFVs. The degree of deviation is calculated as the distance between a space of normative discourses and the deviated discourse. The difference between two norms is also calculated as the difference of their spaces.

Based on the above method, we programed the analyses which consists of following four aspects.

(i) *Local generation space*: The role of the norm is to restrict the generation space into a certain range. For confirming the behavior of norm-based generation, it analyzes the characteristics of the set of normative discourses in each norm.

(ii) *Degree of deviation*: The deviation is a process to transcend the generation from the local generation space at the time. For clarify the actual action of the deviation, it calculates the distance of the deviated discourse from the local generation space.

(iii) *Degree of shift in norm*: For clarifying the manner of actual changes in local generation spaces, it calculates the magnitude of the difference of each norm from the last norm.

(iv) *Novelty of norm*: If local generation spaces are always different with all the past local spaces, the system can produce novel norms and discourses continuously. This analysis calculates the magnitude of the difference of each norm from the most similar norm in all the past norms.

4. Results and Discussions

We executed the system 10000 steps. The input story was same with the story in Ref. 8: A warrior rescues a princess who was abducted by a snake, the plot consisting of 16 total events. The value of n_p, the turning point of the narratee's satisfaction, was 200.

The DFVs of each discourse were automatically calculated. The program preliminary analyzed the range (minimum and maximum), average, and standard deviation of each DFV in all the discourses (Table 1). In addition, 8982 patterns of discourse structures based on DFVs were counted from all the outputs.

Next, the series of discourses was divided into 271 segments, namely norms. The average of segment length (number of discourses) was 36.90 and the minimum and the maximum were 17 and 113.

4.1. *Local generation space*

The analysis (i) analyzed the normative discourses in each norm by the same manner with the above analysis. Table 2 shows two examples of the results. The ranges of DFVs were restricted from the entire set and each norm has different characteristics.

As an issue to be considered, although each local

Table 1. Ranges, averages, and standard deviations of DFVs in all the discourses.

	supplement	complexity	suspense	length	hiding	descriptiveness	repetition	diffuseness	implication	temporal-independency
minimum	0	0	0	11	0	0	0	-9	0	0
maximum	4	50	8	52	10	8	24	20	6	2
average	2.74	8.07	0.69	28.97	3.86	3.44	6.78	2.92	1.90	1.22
SD	1.32	4.48	0.99	6.16	2.34	2.17	3.67	4.21	1.29	0.71

Table 2. The local generation spaces in $norm_1$ and $norm_{181}$.

$Norm_1$ (cycles 1-113)

	supplement	complexity	suspense	length	hiding	descriptiveness	repetition	diffuseness	implication	temporal-independency
minimum	0	0	0	11	3	0	0	-5	0	0
maximum	0	0	0	13	5	0	0	-3	0	0
average	0.00	0.00	0.00	12.30	3.70	0.00	0.00	-3.70	0.00	0.00
SD	0.00	0.00	0.00	0.65	0.65	0.00	0.00	0.65	0.00	0.00

$Norm_{181}$ (cycles 6674-6710)

	supplement	complexity	suspense	length	hiding	descriptiveness	repetition	diffuseness	implication	temporal-independency
minimum	4	8	0	19	5	1	3	-5	1	0
maximum	4	40	5	32	8	3	14	6	3	0
average	4.00	14.94	1.97	25.67	6.72	2.42	6.50	-0.22	2.61	0.00
SD	0.00	7.20	1.26	3.24	0.90	0.72	2.75	2.69	0.54	0.00

generation space had different characteristics, the timing in which the narratee's expectation is saturated is arbitrary defined by the value of n_p. On one hand a lot of overlapped discourses with others were appeared in generation spaces which have small ranges like $norm_1$. On the other hand, large generation spaces like $norm_{181}$ were moved to the next norm before the space was not filled sufficiently. A solution is to redefine the saturation as the filling of the local generation space. This calculation will be embedded into the narratee side.

4.2. *Degree of deviation*

The analysis (ii) calculated the degree of deviation in each norm based on the distance between the local generation space and the deviated discourse. As the result, on one hand about 66% of deviated discourses were positioned outside of the space at the time. On the other hand, about 34% of deviated discourses were included in the space (i.e., the deviation was failed in a practical sense). Such failures occurred due to the partial overlapping between the current local generation space and the deviated (subsequent) space. A solution is to use this analyzing method in the narratee mechanism for judging the success or failure of the deviation.

4.3. *Degree of shift in norm*

The analysis (iii) calculated the degree of shift in each norm based on the difference of each local generation space from the last space. We confirmed that the local generation spaces were gradually shifted. It means that the holistic diversity of generated discourses was arisen through the restriction of generation space based on the norm and the accumulation of small shifts in the norm.

4.4. *Novelty of norm*

The analysis (iv) calculated the novelty of each norm based on the difference of each local generation space from the most similar norm in the past. We clarified that the novelty was gradually decreased through the generation cycle. The reason is that there is a limitation of the possible combinations of the parameters' values in the generative goal and expectation.

5. Conclusion

In this paper, we analyzed a series of discourse structures produced by the narrative discourse system. We used an analyzing program based on a conceptual framework of "norm and deviation". The results quantitatively showed that the diverse discourse structures were arisen through the restriction of generation space based on the norm and the accumulation of small shifts in the norm.

In addition, mainly two issues in the mechanism were clarified: the saturation of the narratee's expectation is arbitrary defined regardless the actual reception of generated discourses, and the narrator fails the deviation often in a practical sense. A solution is to embed the analyzing program into the narratee mechanism for controlling the generation cycle based on the actually generated discourses.

The diversity of generable narratives will an important ability for narrative generation systems. The idea of analyzing method proposed in this paper will be applied to other narrative generation systems for clarifying the holistic generation ability.

References

1. J. P. Rowe, S. W. McQuiggan, J. L. Robison, D. R. Marcey, and J. C. Lester, STORYEVAL: An empirical evaluation framework for narrative generation, in *Papers from the 2009 AAAI Spring Symposium, TR SS-09-06* (2009), pp. 103–110.
2. J. Zhu, Towards a mixed evaluation approach for computational narrative systems, in *Proc. 3rd Int. Conf. Computational Creativity* (2012), pp. 150–154.
3. R. Pérez y Pérez, The three layers evaluation model for computer-generated plots, in *Proc. 5th Int. Conf. Computational Creativity* (2014).
4. F. Peinado, V. Francisco, R. Hervás, and P. Gervás, Assessing the novelty of computer-generated narratives using empirical metrics, *Minds & Machines*, **20**(4) (2010) 565–588.
5. T. Ogata, Expanded literary theory for automatic narrative generation, in *Proc. Joint 7th Int. Conf. Soft Computing and Intelligent Systems and 15th Int. Symp. Advanced Intelligent Systems* (2014), pp. 1558–1563.
6. T. Akimoto and T. Ogata, An information design of narratology: The use of three literary theories in a narrative generation system, *The International Journal of Visual Design*, **7**(3) (2014) 31–61.
7. T. Akimoto and T. Ogata, A narratological approach for narrative discourse: Implementation and evaluation of the system based on Genette and Jauss, in *Proc. 34th Annual Conf. the Cognitive Science Society* (2012), pp. 1272–1277.
8. T. Akimoto and T. Ogata, Towards a discourse mechanism in narrative generation system: Proposal of a system introducing narrative discourse theory and reception theory, *Cognitive Studies*, **20**(4) (2013) 396–420.
9. G. Genette, *Narrative discourse: An essay in method*, transl. J. E. Lewin, (Cornell University Press, NY, 1980).
10. H. R. Jauss, Literary history as a challenge to literary theory, in *Toward an Aesthetic of Reception*, transl. T. Bahti, (University of Minnesota Press, MN, 1982), pp. 3–45.