

ISASE-MAICS 2018

Classifying Personalities of Comic Characters Based on Egograms

Byeongseon PARK**, Kanae IBAYASHI * and Mitsunori MATSUSHITA *

* Faculty of Informatics, Kansai University, 2-1-1, Reizanji-cho, Takatsuki-shi, Osaka 569-1052, Japan
mat@res.kutc.kansai-u.ac.jp

** Graduate School of Informatics, Kansai University, 2-1-1, Reizanji-cho, Takatsuki-shi, Osaka 569-1052, Japan
K281401@kansai-u.ac.jp

Abstract: The aim of this research is to develop a search system for comics based on the personalities of appearing characters. For this purpose, this paper describes the classification of characters using egograms, which are used to classify personalities. In the proposed method, texts that express a comic book character's personality are acquired from web resources, and semantic vectors are allocated based on these texts using egograms. The resulting egogram pattern is used to estimate typical properties. Our experiment reveals that the performance accuracy of this classification method is 55.0%.

Keywords: Personality Classification, TEGII, Web Resource, Comic Computing

1. INTRODUCTION

In Japan, the number of newly published comic book titles is increasing every year, with more than 12,000 titles published in 2017. It is difficult to find comics with content that conforms to the interests of users from this huge number. In order to solve this problem, it is necessary to create an information access support system that can grasp the contents of each comic, and select and present them according to a user's interests[10].

In this research, we aim to achieve such a solution by creating a retrieval system that considers the personality types of appearing characters. Users tend sympathize deeply with characters who embody many of their favorite elements or emotions they identify with. Therefore, comics in which characters with such traits appear are highly likely to match a user's preferences. In addition, characters play an important role in composing the story of a comic. For example, even for comics following the same concept, the story development can differ considerably if the characters' personalities are different. For this reason, in this research we aim to realize a comic search that focuses on the similarity between the character's personality and the user's preferences, with the belief that the attributes of an appearing character form a criterion for selecting comics.

2. DIAGNOSIS OF PERSONALITY BY EGOGRAM

An egogram [1] is a personality diagnosis method based

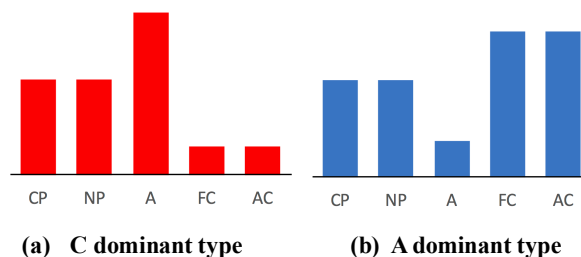


Figure 1: Example of Egogram pattern in TEGII

on the theory of exchange analysis, and it identifies character types using the question paper method. This method is said to have high objectivity, reliability, and validity, and is utilized in various situations such as educational and medical approach [4-9].

Currently, the egogram that is most widely employed in Japan is the second version of the Tokyo University Egogram (hereafter referred to as TEGII) [3]. In TEGII, we measure the character of a person divided into the following five ego-state areas.

- Critical Parent (CP): Mind-concerning aspects such as obligation, sense of responsibility, belief, and pursuit of ideals.
- Nurturing Parent (NP): Mind-concerning aspects such as acceptance, compassion, development of self, and other mind and body aspects.
- Adult (A): Mind-concerning aspects such as logicity, rationality, objectivity, and plannability.
- Free Child (FC): Mind-concerning aspects such as autonomy, activity, aggressiveness, and creativity.

Table 1: Example of words acquired from the web

Gong Freaks (ゴンフリークス)	Uzumaki Naruto (うずまきナルト)	Son Goku (孫悟空)
bright(明るい)	never give up (諦めない)	bright (明るい)
full of curiosity (好奇心旺盛)	Justice (正義感)	Fool (お人好し)
honest (裏表がない)	self-indulgence (我が儘)	honest (裏表がない)
honest(素直)	reckless (無鉄砲)	overconfidence (自信過剰)
calm(冷静)	straight (まっすぐ)	honest(素直)
sore loser (負けず嫌い)	sore loser (負けず嫌い)	friendly (優しい)
friendly (優しい)	friendly (優しい)	calm(冷静)

- Adapted Child (AC): Mind-concerning aspects such as cooperativeness, social nature, compromise, and dependence.

TEGII classifies the personality into 29 types according to the degrees of correspondence to these five areas. For example, in the left diagram of Fig. 1, focusing on the fact that FC and AC are of similar height, the personality is classified as "C dominant type." This shows that even if the subject is an adult, they exhibit a childlike selfish state, and behave according. Furthermore, the right diagram of Fig. 1, focuses on the point that only A protrudes, and is classified as "A dominant type." This indicates that the subject exhibits many logical, intellectual, and planned actions.

In this research, we propose a method to classify personalities of comic book characters according to TEGII.

3. PERSONALITY CLASSIFICATION OF A COMIC BOOK CHARACTER

3.1 Proposed Method

In the proposed method, we first extract sentences that express briefly a character's personality from document data that can be collected from the web, and create a personality word dictionary for the personality classification. In our research, personality words refer to words (e.g., bright, calm, friendly) and short sentences (e.g., caring person, deep affection) that briefly express personality.

In order to judge the personality type of a character, we create a vector (hereafter referred to as a personality vector) by referring to the personality words obtained

Table 2: Examples of personality features

Egotic condition	value is high	value is low
CP	mandatory	easygoing
	responsibility	friendly
	stubborn	loose
NP	caring person	nervous
	kind	own pace
	meddle	frank
A	rational	naive
	calm	fool
	cold	not realistic
FC	curiosity	quiet
	free-roaring	meek
	bright	moderate
AC	cooperativeness	lively
	honest	positive
	obedience	own pace

from the resources on the web and the created dictionary. The degree of similarity between this personality vector and the preregistered basic egogram pattern is calculated, and those closest in shape are taken as the personality type of the character.

3.2 Personality Word Dictionary Creation

First, we collected descriptions expressing character 's personalities from Wikipedia¹, Niconico Pedia²(ニコニコ大百科), pixiv encyclopedia³(pixiv 百科事典), and Aniota wiki⁴(アニヲタ wiki), and created a dictionary, which represents the cornerstone of each character' s personality classification.

The basic vocabulary for creating a dictionary targets those words used in relation to each of the main 100 characters appearing in the top 19 comic books (ranked in June 2017) published in the manga chronological issue number stated on the Mangazenkan.com⁵(漫画全巻ドットコム). From the document collected from the above website, the words and short sentences that briefly express personalities are manually handled as personality words for the corresponding character. Table 1 shows examples of extracted personality words.

We created a personality word dictionary based on the personality words that were collected and extracted as a result of the above process. Each of the five regions of the egogram described in TEGII represents a different personality, and various personality features can be read

¹ <https://ja.wikipedia.org/wiki/>

² <http://dic.nicovideo.jp/>

³ <https://dic.pixiv.net/>

⁴ <https://www49.atwiki.jp/aniwotawiki/>

⁵ <https://www.mangazenkan.com/>

Table 3: Example Personality Word dictionary

Word	Area of Egogram				
	CP	NP	A	FC	AC
calm	0	0	1	0	0
mandatory	1	0	0	0	0
easygoing	-1	0	0	0	0
kind	0	1	0	0	0
fool	0	1	-1	0	0
own pace	0	-1	0	0	-1
free-roaring	0	0	0	1	0
honest	0	0	0	0	1

from the high and low values. For example, a high value of CP indicates personality features such as "mandatory (義理堅い)" or "stubborn (頑固)," and a low value of CP indicates character characteristics such as "easygoing (おっとり)" or "loose (ルーズ)." Table 2 shows examples of personality features. In this paper, the vector created in this manner is called the personality vector.

With respect to these 664 personality words, an evaluation is performed with reference to the vocabulary described in TEGII as to whether there is a match with personality traits of each region of TEGII, and a vector based on the five regions (CP, NP, A, FC, and AC) of TEG II is created. At this time, values of "1" and "-1" are assigned when the values of a region are high and low, respectively, with "0" corresponding to not matching at all. For example, because the personality word "stubborn (頑固)" is a feature that appears when the CP value is high, the personality vector corresponding to "stubborn (頑固)" is (1, 0, 0, 0, 0), where CP is 1. Because there are not many personality features described in TEGII, we have expanded it to include words within two hops of similar words recorded in weblio zisyo⁶. Through this processing method, 534 words out of the 664 used to create the dictionary could be given vectors. A part of the result is shown in Table 3.

3.3 Method of Estimating Personality Types of Comic Characters

We propose a method of judging the personality of each character with reference to the personality word dictionary created in 3.2. In the proposed method, we characterize a comic character's personality type by examining the personality words from the personality word dictionary contained in web descriptions related the character.

Next, a personality vector is generated by referring to the personality word dictionary. The personality vector is

Table 4: Personality words of "Monkey D Luffy"

Responsibility (責任感)	King (王様)	Leadership (リーダーシップ)
Charisma (カリスマ)	Daring (豪胆)	Tolerance (包容力)
big heart (器が大きい)	Humanity (人情)	Friendly (優しい)
Flexible (柔軟)	Argumentative (理屈)	free-roaring (自由奔放)
Spontaneously (行動力)	self-oriented (自分本位)	Renewal (根明)
Fierce (凶暴)	Reckless (無鉄砲)	Rough (荒い)
Reckless (無謀)	Carefree (能天気)	Optimism (楽観主義)
Simple (単純)	Innocence (無邪気)	Serious (真面目)
Pure (純粹)	Strict (厳しい)	Dignity (威厳)
Straight (ストレート)	Capricious (気まぐれ)	Messenger (お調子者)
never give up (諦めない)	Insensitivity (鈍感)	easy to be fooled (騙されやすい)
Bright (明るい)	Magnificent (豪快)	

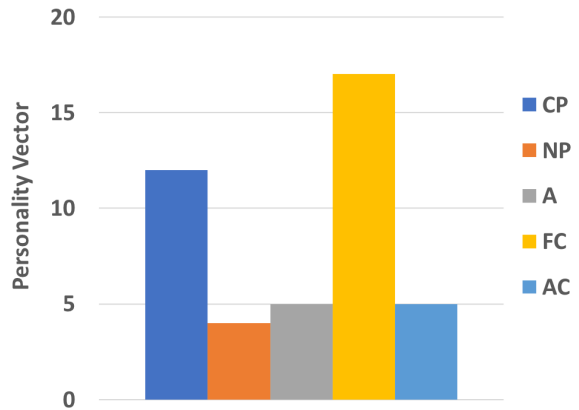


Figure 2: Egogram of "Monkey D Luffy"

obtained as a linear sum of the personality vectors of each personality word. Because the number of personality words obtained for each comic character is different, an adjustment is made such that the total numbers are the same among characters.

Based on the value of the personality vector obtained for each comic character and the value of the basic type of the egogram, the similarity f was calculated according to Equation (1) to estimate the personality of the a character.

Equation (1) measures the distance between the personality vector $(x_1, x_2, x_3, x_4, x_5)$ of each comic character and the basic type $(y_1, y_2, y_3, y_4, y_5)$ of the egogram. The closer the calculated value is to 0, the higher the similarity.

$$f = \sum_{i=1}^5 (x_i - y_i)^2 \quad (1)$$

⁶ <https://www.weblio.jp/>

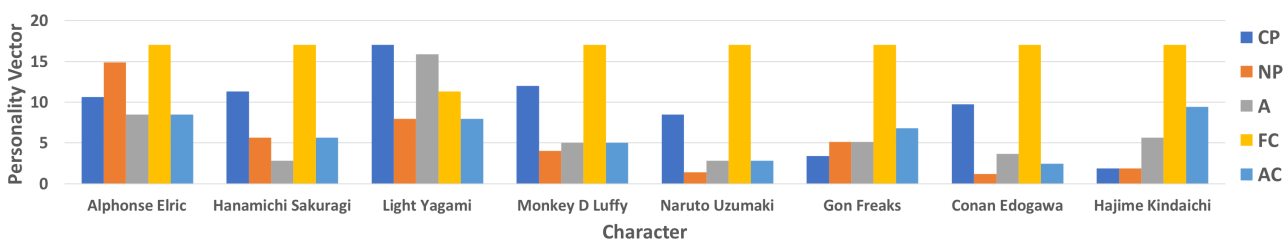


Figure 3: Personality Vector obtained by proposed system

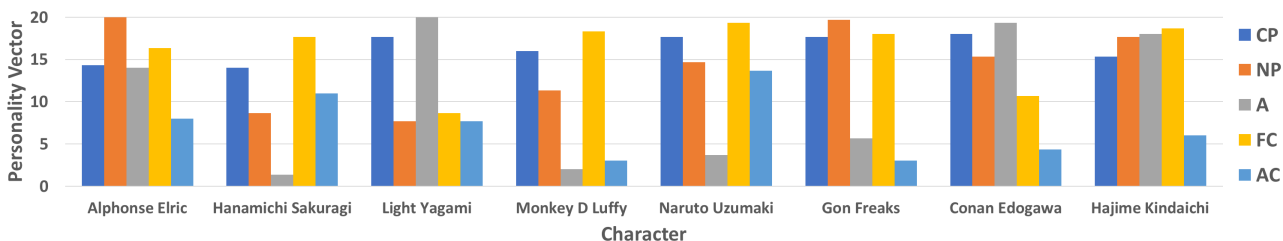


Figure 4: Personality Vector obtained by experiment

For example, for "Monkey · D · Luffy", the hero of "One Piece" (written by Eiichiro Oda, Shueisha), the personality words shown in Table 4 were extracted.

In this case, the personality vector is calculated based on these words (12, 4, 5, 17, 5) (see Figure 2). When this personality vector is applied to Equation 1, it becomes "inverse N type 3," "reverse N type 2," "FC dominant" in descending order of the degree of agreement with the basic types of the egogram, so that the personality type of the character is determined to be "reverse N type 3." "Reverse N-type 3" has the characteristics of self-centered, selfish, intimidated, emotional, and so on, and can be said to be valid for the personality of this character.

4. EXPERIMENT

4.1 Experimental Procedure

We conducted a user experiment to verify whether the personalities of comic characters extracted using the proposed method match the personalities recognized by users.

First, we asked the users to create an egogram of characters using the question paper method. The subject data of the experiment consisted of 20 characters from 17 comics, and data was collected from three experimental collaborators per character.

Fifteen people (nine men and men women) aged from their 20s to 40s participated as experiment collaborators in this experiment. Because one cannot provide answers without being familiar with a character, it was necessary to (1) read the comic in which the character to be

evaluated appears, and (2) recognize the character to be evaluated. As a condition, we decided which characters would be assigned each experimental collaborator. Under this condition, between one and 18 characters were assigned to each collaborator. The reason for the different number of assignments is that the number of comics that had been read varied greatly for each collaborator.

In the experiment, we asked how the assigned comic characters corresponded to the questions (e.g., "is it easy to be influenced by what others say?" or "has a sense of humor?") of all 50 egograms described in TEGII. In this case, we calculated the total value following TEGII, with a scale of two points if applicable to the question, zero points if not applicable, one point if not sure.

We compared the judgment results for the top three using the proposed method, with the type indicated by the judgment result of three experimental collaborators and verified by the coincidence ratio. At this time, if the types indicated by the judgment results of three experimental collaborators included in the top three personality types extracted using the proposed method, this was regarded as being consistent.

4.2 Experimental Result

The results for the proposed system and the top three types resulting from the questionnaire assigned 55.0% of the comic characters with the same types (11 out of 20 characters). Figure 3 shows the personality vectors for each character calculated using the proposed method, and Figure 4 shows the judgment results from the experiment partners. Table 5 shows a part of the results of comparing

Table 5: Experiment results (matched personality types are bold)

Character name	Similarity	Result of system			Result of experiment collaborators
		1st	2nd	3rd	
Alphonse Elric	0.057	Flat type 1	Trapezoid type 1	P dominant type	Flat type 1
Hanamichi Sakuragi	0.081	Reverse N type 3	Reverse N type 2	U type 2	Reverse N type 3
Light Yagami	0.086	Reverse N type 1	Flat type 1	P dominant type	Reverse N type 1
Monkey D Luffy	0.140	Reverse N type 3	Reverse N type 2	FC dominant type	Reverse N type 3
Tatsuya Uesugi	0.145	N type 1	C dominant type	Flat type 1	N type 1
Kankichi Ryotsu	0.171	Reverse N type 3	FC dominant type	Reverse N type 2	Reverse N type 3
Gintoki Sakata	0.263	M type	FC dominant type	Reverse N type 3	Reverse N type 3
Kurosaki Ichigo	0.286	NP dominant type	P dominant type	Flat type 2	P dominant type
Zaodyeek Killua	0.323	Flat type 3	FC dominant type	C dominant type	W type
Souma Yukihira	0.366	FC dominant type	Reverse N type 2	Reverse N type 3	Reverse N type 3
Edward Elric	0.373	FC dominant type	Flat type 3	Flat type 1	Reverse N type 3
Roronoa Zoro	0.494	Reverse N type 1	Flat type 3	A dominant type	AC recessive type
Syouyou Hinata	0.526	U type 2	Reverse N type 2	Reverse N type 3	Reverse N type 3
Ellen Yeager	0.561	Reverse N type 2	Reverse N type 3	FC dominant type	Reverse N type 1
Son Goku	0.647	FC dominant type	Flat type 3	C dominant type	Reverse N type 3
Kenshin Himura	0.664	M type	NP dominant type	Flat type 1	P dominant type
Naruto Uzumaki	0.837	Reverse N type 2	Reverse N type 3	FC dominant type	Reverse N type 3
Gon Freaks	1.006	FC dominant type	C dominant type	Flat type 3	Reverse N type 3
Conan Edogawa	1.380	Reverse N type 2	Reverse N type 3	FC dominant type	P dominant type
Hajime Kindaichi	1.416	AC dominant type	C dominant type	FC dominant type	Flat type 1

the types of egogram.

In the experimental results, the highest degree of coincidence between the system and users was "Alphonse Elric" (0.057), who is a character from "Full Metal Alchemist (鋼の錬金術師)" (written by Hirohiro Arakawa, Square Enix). The lowest was "Hajime Kindaichi" (1.416), who is a character from "Kindaichi Boy's Case Files (金田一少年の事件簿)" (written by Sato Fumiya, Kodansha).

4. DISCUSSION

The accuracy of the proposed method was shown by the experiment to be 55.0%. Because the ratio corresponding to three of the 29 personality types would be 10.3% under random conditions, the estimation of the personality type using the proposed method seems to have demonstrated an effect. From this, it is suggested that the egogram method used in this paper can be applied not only to the personalities of humans, but also of comic characters.

On the other hand, the reasons why the classification accuracy remained at 55.0% are as follows. The first factor is the extraction of unnecessary words as personality words of comic characters. In this approach, a method of extracting personality words matching with words registered in a personality word dictionary was adopted. Here, it is not possible to consider the meanings of words by context. For example, it is desirable that only the words "bright" and "energetic" are extracted from the sentence "rather than calm, bright, and energetic," but "calm" is also extracted using the current method. In addition, although the sentence "severe person" represents a "severe personality," "severe reality" does not indicate that the personality is severe. By wrongly extracting these words, the classification accuracy is deteriorated.

The second factor is that personality words cannot be extracted according to style. For example, from sentences such as "I care about my friends more than anyone, I care very much" and "I have a strong commitment to protect important people," traits such as "gentle," "friendly," and

"passionate" can be estimated. However, because the proposed method extracts words and short sentences through superficial matching, personality words cannot be extracted from these sentences, and the intended personality cannot be estimated.

The third factor is that as a story evolves, the personality can change. This is a tendency found in many long-term serial works, and sometimes there are differences in personality depending on the selected time period. In comics, characters are often subject to extreme characterization, and there are cases where characters with a gentle personality change to having a cruel personality because of certain events. An egogram cannot capture such personality changes due to narrative features.

In the future, we will consider these influences and improve the accuracy of the method.

5. CONCLUSION

In this paper, in order to construct a comic book search support system, a focus was given to the personalities of characters, which considerably affects the story. A method was proposed to classify the personalities of characters based on egograms. Experimental results confirmed that the classification of personalities using egograms is possible. Thus, because the similarities between the personalities of characters and a search request is to be calculated, results such as "a comic book character with a personality that is similar to this character appears" are expected. In the future, we will extend this research by expanding the personality word dictionary and improving the accuracy of the personality type estimation.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 15K12103.

REFERENCES

- [1] John M. Dusay, M.D.; Egograms and the constancy hypothesis, *Transactional Analysis Journal*, Vol.2 No.3, pp.37-41, 1972.
- [2] Shunji Hirose, Manabu Tatebayashi, Kumi Saito, and Satoru Yoshikawa; A study of psychological factors of bruxism: Analysis of risk factors, *Japanese Society of Psychosomatic Medicine*, Vol.55, No.7, pp.864-872, 2015. (In Japanese)
- [3] Michiko Wada, Asami Watanabe, Miho Ichimura and Yutaka Matsui; New egogram developed by web survey, *Tsukuba Psychological Research*, Vol.53, pp.63-71, 2017. (In Japanese)
- [4] Liang Cai, Sakae Kaneko and Eishin Morita; Changes in salivary chromogranin A levels in adults with atopic dermatitis are correlated with changes in their condition, *The Journal of Dermatology*, 2018.
- [5] Min Jung; Influence of ego states, self-esteem, and empathies on interpersonal relationship of nursing students, *Journal of the Korea Academia-Industrial cooperation Society*, Vol.16, No.7, pp.4614-4620, 2015.
- [6] Gao Xu, and Minoru Uehara; Design of a sports mental cloud, *Advanced Information Networking and Applications Workshops*, 31st International Conference on. IEEE, pp.443-448, 2017.
- [7] Ng Annalyn, Maarten W. Bos, Leonid Sigal, and Boyang Li; Predicting personality from book preferences with user-generated content labels, *IEEE Transactions on Affective Computing*, 2018.
- [8] Teruyuki Iijima and Shin Tezuka and Ryuya Uda; Method for Extraction of characteristics of personal characters from life log, *Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication*, pp.110:1-110:8, 2014.
- [9] Margaret L. Kern, Johannes C. Eichstaedt, H. Andrew Schwartz, Lukasz Dziurzynski, Lyle H. Ungar, David J. Stillwell, Michal Kosinski, Stephanie M. Ramones, and Martin E. P. Seligman; The online social self: An open vocabulary approach to personality, *Assessment*, Vol.21-2, pp.158-169, 2014.
- [10] Byeongseon Park, Kahori Okamoto, Ryo Yamashita, and Mitsunori Matsushita; Designing a comic exploration system using a hierarchical topic classification of reviews, *Information Engineering Express International Institute of Applied Informatics*, Vol.3, No.2, pp.45-57, 2017.