

A support interface for remembering events in novels by visualizing time-series information of characters and their existing places

Yoko Nishihara, Jiaxiu Ma, and Ryosuke Yamanishi*

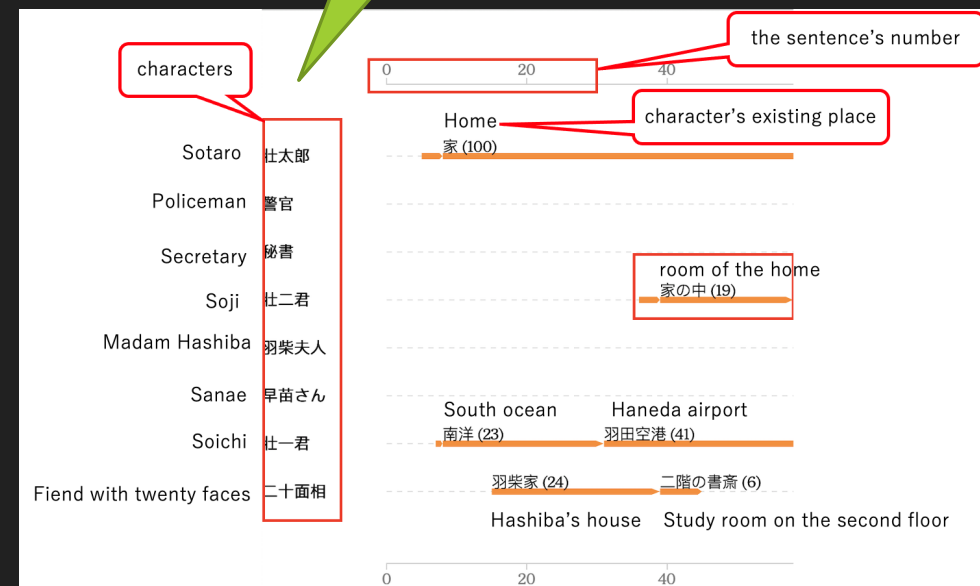
Ritsumeikan University, Japan

*Kansai University, Japan

What we proposed

- We proposed a support interface for remembering events in novels by visualizing time-series information of characters and their existing places.
- This interface supports remembering the events in the already read parts.
- This interface may remove rereads and help users in resuming reading novels.

Horizontal axis: the sentence number
Vertical axis: character names



Research background

- People use electronic devices to read e-books.
- It is expected that the style of reading several e-books in parallel will spread widely.
- **If there is a time gap between readings, people may forget what has happened in the already read parts of each book.**
 - Especially if many characters appear in a novel.
- The more times people reread books, the more difficult they enjoy reading.



Objective of this research

- We propose a support interface for remembering events in novels by visualizing time-series information of characters and their existing places.
- The target users are people who read **several novels in parallel**.
- We believe that the visualized information can support the users to remember when and where each character did something and what he/she thought in the already read parts.

Outline of the proposed interface

(a) Displaying novel titles and texts.

The left-side menu shows the list of novels.

The right-side window shows the texts of the selected novel.

STORYS

- 怪人二十面相
- 少年探偵団
- 少年探偵長
- 銀河鉄道の夜

怪人二十面相

江戸川乱歩

【テキスト中に現れる記号について】

《 》：ルビ
(例) 盗賊《とうぞく》

| : ルビの付く文字列の始まりを特定する記号
(例) 一 | 騎《き》

[#] : 入力者注 主に外字の説明や、傍点の位置の指定
(例) はしがき [# 「はしがき」は中見出し]

はしがき [# 「はしがき」は中見出し]

そのころ、東京中の町という町、家という家では、ふたり以上の人が顔をあわせさえすれば、まるでお天気のあいざつでもするように、怪人「二十面相」のうわさをしていました。「二十面相」というのは、毎日毎日、新聞記事にぎわしている、ふしぎな盗賊《とうぞく》のあだ名です。その賊は二十のまったくちがった顔を持っているといわれていました。つまり、変装《へんそう》がとびきりじょうずなのです。どんなに明るく場所、どんなに近づいてながめても、少しも変装とはわからない、まるでちがった人に見えるのだそうです。老人にも若者にも、富豪《ふごう》にも乞食《こじき》にも、学者にも無頼漢《ぶらいかん》にも、いや、女にさえも、まったくその人になりきってしまうことができるといいます。では、その賊のほんとうの年はいくつで、どんな顔をしているのかというと、それは、だれひとり見たことがありません。二十種類の顔を持っているけれど、そのうちの、どれがほんとうの顔なのか、だれも知らない。いや、賊自身でも、ほんとうの顔をわすれてしまっているのかもしれない。それほど、たえずちがった顔、ちがった姿で、人の前にあらわれるのです。

switch the display to remember the events

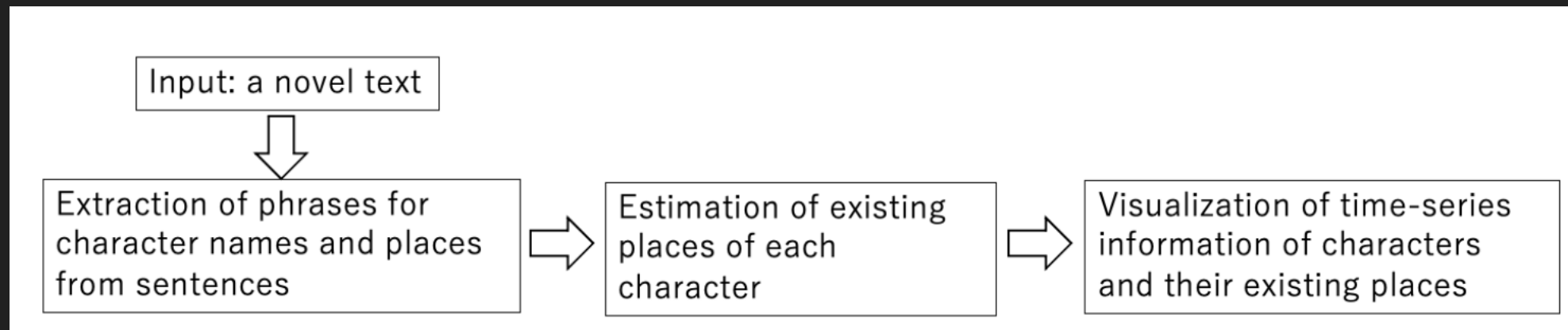
(b) Visualizing time-series information of character and their places.



restart reading

Procedures to obtain the visualization

1. Extraction of phrases for characters names and places
2. Estimation of existing places for each character
3. Visualizing of time-series information of characters and their existing places



1. Extraction of phrases for characters names and places

○ **CRFs** as the named-entity recognition method is used.

1. Each sentence in a text is divided into phrases with the information of part of speeches by a morphological analyzer (MeCab+NEologd).

2. An annotator gives BIOES tags for character names and places on phrases to make a dataset.

3. The dataset is learned by CRFs to obtain an extraction machine of phrases for character names and places.

Tag	Definition
B-CHAR	The beginning of a phrase for a character name
I-CHAR	The intermediate of a phrase for a character name
B-POS	The beginning of a phrase for a place
I-POS	The intermediate of a phrase for a place
O	Others

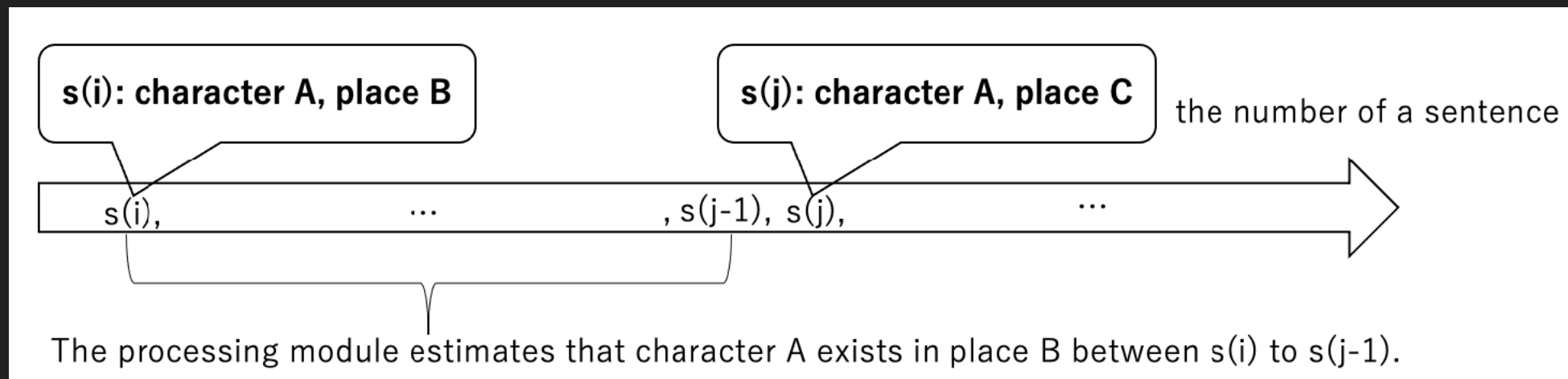
2. Estimation of existing places for each character (1/2)

- We assume **if one sentence includes phrases for both a character name and a place**, the character name should be related to the place, **which means the character may exist in the place**.
- We conducted preliminary experiments to show the validity of the assumption.
- We found 71% sentences satisfied the assumption.

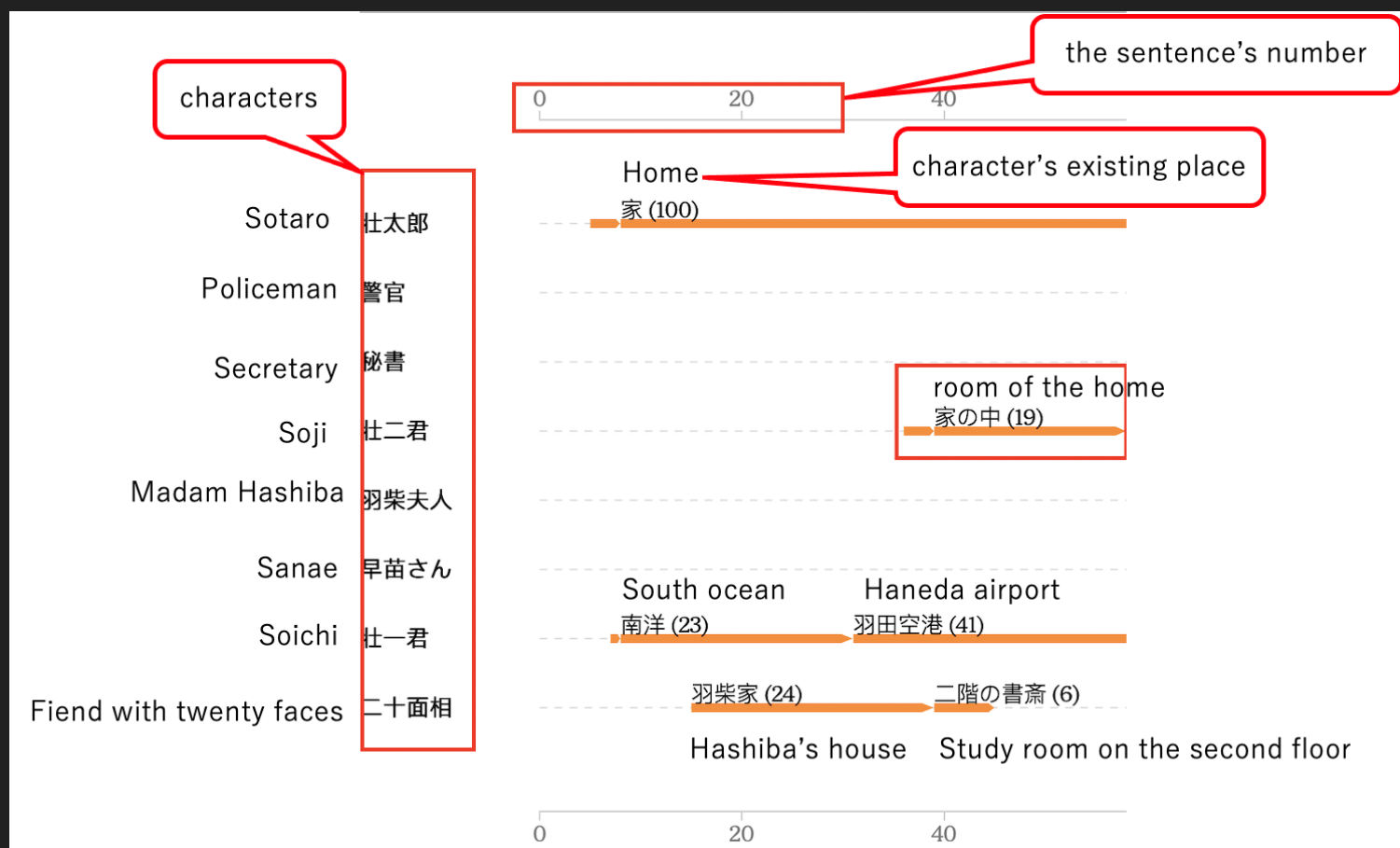
Novel	Only character name	Only place	Both	Neither	Total
Momo Taro	34	12	12	24	126
Little Red Riding Hood	53	17	32	24	82
Night on the Galactic Railroad	251	44	55	274	624
Total	338	73	99	322	832

2. Estimation of existing places for each character (2/2)

- The characters in novels would stay in the same place for a certain period.
- If a sentence $s(j-1)$ include only a phrase for a character name A, his/her existing place is estimated by the sentence $s(i)$ that includes both phrases for the character name A and a place B ($i < j$).



3. Visualizing of time-series information of characters and their existing places



The user can remember what each character did/thought something.

The user also can recognize where each character exits at a certain time.

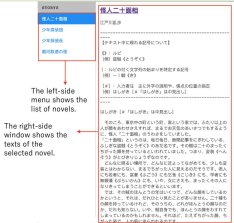
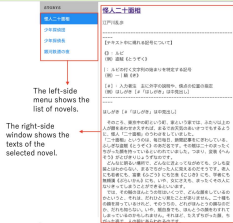
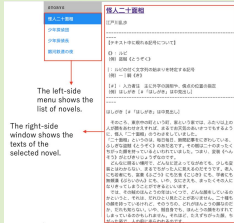

Evaluation experiments (1/2)

- Experimental procedures
 - 1. Each participant **reads four novels in a day** with an assigned interface. He/she reads a part of the novel.
 - 2. After one week, he/she **answers quizzes** about events in the novels.
 - 3. The experimenter compares the correct rates of the quizzes among groups.



Evaluation experiments (1/2)

○ 21 participants and 3 groups.

Assigned interface	Group A (7 participants)	Group B (7 participants)	Group C (7 participants)
When reading novels	 <p>The left-side menu shows the list of novels.</p> <p>The right-side window shows the texts of the selected novel.</p> <p>part(a) of the interface</p>	 <p>The left-side menu shows the list of novels.</p> <p>The right-side window shows the texts of the selected novel.</p> <p>part(a) of the interface</p>	 <p>The left-side menu shows the list of novels.</p> <p>The right-side window shows the texts of the selected novel.</p> <p>part(a) of the interface</p>
When answering quizzes	<p>They did not use any references</p>	<p>They reread novels in 3 minutes.</p>	 <p>part(b) of the interface</p>

Used novels and desinged quizzes

- We chose 4 long novels from AOZORA BUNKO.
- We asked about 8000 Japanese alphabets to read.
 - We assume one reading is for 20 minutes and 400 Japanese alphabets.
- Quizzes were about characters and places.
 - The participants answered by writing words and choosing options.

Title	Author	# of Japanese alphabets to read
Boy Detectives Club	Ranpo Edogawa	9,979
Shonen Tantei Cho	Jyuza Unno	6,101
Fiend with Twenty Faces	Ranpo Edogawa	8,907
Night on the Galactic Railroad	Kenji Miyazawa	8,903

Quiz type	Number of quizzes	Example of quizzes (Novel title)
A character	9(5)	What did Giovanni buy before leaving? (Night on the Galactic Railroad)
Characters	3(1)	Who did clime up Kan-nuki mountain with Haruki? (Shonen Tantei Cho)
A place	5(0)	Where was Hashiba's jewel? (Fiend with Twenty Faces)
Places	3(2)	Where did Kei run to bright town from? (Boy Detectives Club)





Actually we believe the design of experiment to evaluate the interface as the remember is so interesting and may be applied to other topics too.

Experimental results

Averaged correct rates of the quizzes

- The correct rate of Group C was the highest.
- The difference between Group A and B was +10%.
- The difference Group A and C was + 28%.
- The visualization of time-series information could support more for remembering the events than rereading.

Group	Correct rates
A	26%
B	36% (+10%)
C	54% (+28%)

Assigned interface	Group A (7 participants)	Group B (7 participants)	Group C (7 participants)
When reading novels	 part(a) of the interface	 part(a) of the interface	 part(a) of the interface
When answering quizzes	They did not use any references	They reread novels in 3 minutes.	 part(b) of the interface

Discussion of the correct rates of the quizzes about a character(s)

Group	A character	Characters	A place	Places
A	40%	14%	6%	33%
B	38% (-2%)	43% (+29%)	20% (+14%)	52% (+19%)
C	57% (+17%)	62% (+48%)	40% (+34%)	62% (+29%)
Average	45%	40%	22%	49%

- The correct rate of characters was lower than that of a character.
- The complexity of events **about characters** may cause the lower correct rate.
- Rereading and the proposed interface supported participants to remember the events, so the higher correct rates were obtained.

Discussion of the correct rates of the quizzes about a place(s)

Group	A character	Characters	A place	Places
A	40%	14%	6%	33%
B	38% (-2%)	43% (+29%)	20% (+14%)	52% (+19%)
C	57% (+17%)	62% (+48%)	40% (+34%)	62% (+29%)
Average	45%	40%	22%	49%

- The correct rate of a place was lower than that of places.
 - The quizzes about a place were all descriptive type.
- The correct rate of all descriptive type quizzes was 40% in Group C.
- The result might indicate that the participants with the proposed interface could remember 40% events at the highest.
- We improve the quizzes to evaluate the proposed interface more precisely.

Conclusion

- This paper proposed a support interface for remembering events in novels by visualizing time-series information of characters and their existing places.
- Evaluation experiments verified that the proposed interface could support participants more significantly than rereading.
- We improve the quizzes to evaluate the proposed interface and find the limitation of the interface.