Towards Smarter Democracy: An Agent-based Large-scale Discussion Support System

Takayuki Ito[1]*, Shota Suzuki[1] Naoko Yamaguchi[1] Tomohiro Nishida[1]
Kentaro Hiraishi[1] and Kai Yoshino[1]
[1]Nagoya Institute of Technology
Japan
{ito,yamaguchi.naoko,nishida.tomohiro,}@nitech.ac.jp,
{suzuki.shota,hiraishi.kentaro,yoshino.kai}@itolab.nitech.ac.jp

Abstract

Discussion is essential for democracy. Providing good support is critical for establishing and maintaining coherent discussions. Large-scale online discussion platforms are receiving great attention as potential next-generation methods for smart democratic citizen platforms. Such platforms require support functions that can efficiently achieve a consensus, reasonably integrate ideas, and discourage flaming. We are developing several crowd-scale discussion platforms and conducting social experiments. In the initial version for such system, we employed **human facilitators** in order to achieve good disucssion. However, we clarified the critical problems faced by human facilitators caused by the difficulty of facilitating crowd-scale online discussions. In this work, we propose an **automated facilitation agent**[Ito and Shiramatsu, 2018] to manage online discussions. An automated facilitator agent extracts the discussion structure from the texts posted in discussions by people. In this paper, we present our current implementation of a crowd-scale discussion support system based on an automated facilitation agent, which extracts discussion structures from text discussions, analyzes them, and posts facilitation messages. We conducted a large-scale social experiment with Nagoya City ' s local government in which our automated facilitation agent worked well.

1 Introduction

Recent developments of AI and information technology have great potential to make our social system smarter. All forms of human social mechanisms, including democracy, market, etc., would have potential to evolve by these developments. One of the most important social mechanisms is **democracy**. For good democacy, large-scale discussion is essential. Large-scale online discussion platforms are receiving great attention as potential next-generation methods for smart democratic citizen platforms [Malone and Klein, 2007; Malone, 2018]. Such platforms require support functions that can efficiently achieve a consensus, reasonably integrate ideas, and discourage flaming. We are developing several crowd-scale discussion platforms and conducting social experiments with private citizens. The first version was called COLLAGREE [Sengoku *et al.*, 2016; Ito *et al.*, 2015; Ito *et al.*, 2014; Ito, 2018], which we employed for a large-scale experiment with Nagoya City, Japan. In this 2013 experiment, we collaborated with its local government and gathered opinions from the public about a next-generation comprehensive plan. Our experiment ran for a two-week period in December with nine expert facilitators. We gathered 266 registered participants, 1,151 opinions, 3,072 visits, and 18,466 views. Ours was first trial where **human facilitators** promoted crowd-scale online discussions.

After the above experiment, we conducted more than 30 experiments [KAWASE *et al.*, 2018; Nishida *et al.*, 2018; Nishida *et al.*, 2017] and clarified the critical problems faced by human facilitators caused by the difficulty of facilitating crowd-scale online discussions. Such discussions often have over a thousand opinions that are posted simultaneously. Many discussion threads become tangled with overlapping opinions. Such elements are characteristic problems for online discussions that are not seen in ordinary face-to-face discussion workshops.

In this work, we propose an **automated facilitation agent**[Ito and Shiramatsu, 2018] to manage online discussions.

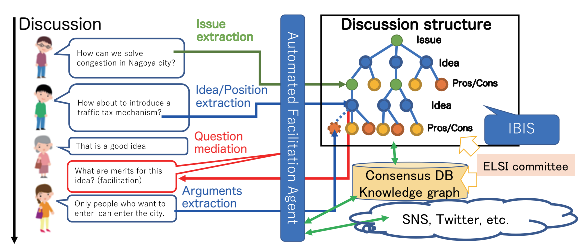


Figure 1: Outline of our system

Figure 1 outlines our system. An automated facilitator agent extracts the discussion structure from the texts posted in discussions by people. The discussion struc-

*Contact Author

ture represents a discussion's semantics. We adopted IBIS structure [Kunz and Rittel, 1970] as a discussion framework because our aim is to create discussions through which people can clarify issues, ideas, and debate merits/demerits. IBIS effectively constructs such discussions. Extending any form of argumentative structure is also easy [Lawrence and Reed, 2017]. Based on the extracted structure, facilitation agents post facilitation messages about the discussion, manage a knowledge database that contains the previous discussion structure, and collect data from other social media.

Figure **??** shows a typical user-interface employed by both facilitators and participants. The following are its typical functions:

- Posting form: users post their opnions from the posting form.
- Discussion board: users can see their posted opinions (postings) at here.
- Theme: the main theme for this discussion is shown
- Points: we provide discussion posints[Ito *et al.*, 2015] as a virtual point for discussion incentive.
- Keyword cloud: Keyword cloud is highlighted so that facilitators and userscan quickly understand what words are being focused on and which are important.

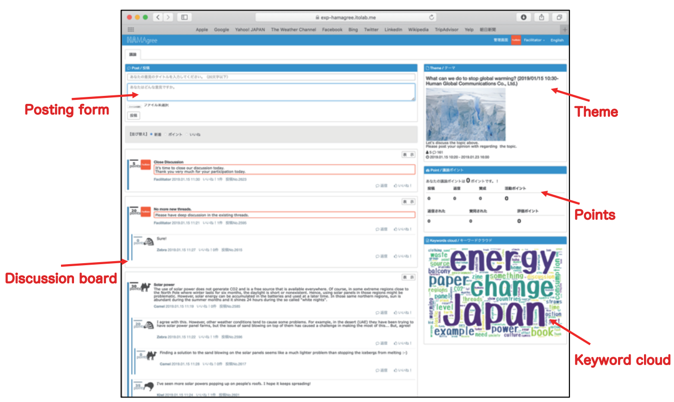


Figure 2: User interface

2 Automated Facilitation Agent

We developed automated facilitation agent software that observes the posted texts, extracts their semantic discussion structures, generates facilitation messages, and posts them to the discussion system. The software also filters inappropriate posts.

The facilitator agent consists of two parts: a discussion structure extraction/visualization mechanism and an observing and posting mechanism. To extract the discussion structure, we utilize deep-learning technologies including BiLSTM [Lample *et al.*, 2016], which first captures meaningful sentences and then important words that are IBIS components: issues, ideas, pros, and cons. After that, it identifies the relations among these IBIS components and unifies these relations and components into one discussion structure.

Extracting a discussion/argumentation structure has been widely studied in the argumentation mining field [Lawrence and Reed, 2017; Stab and Gurevych, 2017; Stab and Gurevych, 2014b; Stab and Gurevych, 2014a]. The main difference between our approach and the argumentation mining field is that ours is based on IBIS structure, which focuses on facilitation. The IBIS structure includes an issue component, which is different from ordinal argumentation structures. Almost all argumentation mining researches use structures that consist of claims, supports, premises, and so on. They do not have issue components in their structure. Issue components are critical for facilitation and innovative discussions. The F scores of extracting issue components exceed 0.80, and the precision score of the identifying links among components are around 0.88. The F score is the harmonic average of the precision and recall. These scores are higher than state-of-the-art argumentation mining, and these results greatly depend on our clearning and annotation efforts on a large discussion dataset that is comprised of over 38 pieces of actual discussion data. After carefully defining our annotation scheme, our annotation results had a kappa value of around 0.66 [Yamaguchi *et al.*, 2018]. kappa value is a statistic which measures inter-rater agreement for qualitative items. The following are the detailed numbers:

- Annotated sentences : 4,972
- Average discussion time : 7.81 hors (15min to 144 hours)
- Average participants : 13 people (4-114 people)
- Average number of postings: 101.8 per a discussion (including facilitation posting)
- Themes: education, disaster mitigation, environments, sightseeing, etc.

By using the extracted structure, the observing and posting mechanism posts facilitation messages. It has around 200 facilitation rules, which have been carefully collated after consultation with professional facilitators. By matching the rules and the obtained structure, facilitation messages are basically generated. These facilitation mechanisms have been implemented by the AWS lambda function and AWS CloudWatch. Our current entire system operates on Amazon Web Services.

Our agent has a fanction that can filter impertinent postings using supervised learning with distributed representation of posted documents and vectors as features. This is because in real social experiments, there could be many harmful contents such as unrelated spam in these discussion platforms, and violent remarks that insult and discriminate against opponents. As result, it becomes necessary to build a discussion platform that allows online users to participate safely by removing inappropriate remarks. To remove inappropriate remarks, understanding and classifying the meanings of documents is needed. To this end, we adopt doc2vec and ELMo to word embedding documents. In addition, we constructed a vectorized document by using document similarity calculation and deep neural networks (Bi-DNN). Table **??** shows the resulting F-values using Bi-RNN. The filtering function itself has higher accuracy.

Method	F-measure
Doc2Vec Ensemble	0.936
Bi-LSTM	0.919
Bi-GRU	0.9164

Figure 3 presents the whole architecture of our system and its user interface. This is done by AWS lambda and CloudWatch, which are scalable enough even if we have many numbers of discussions. We can use English and Japanese as languages.

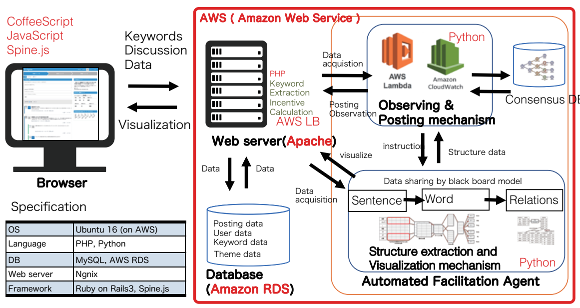


Figure 3: System Architecture

3 Societal Experiment with Nagoya Local Government

We conducted a real world experiment with the Nagoya municipal government from November 1 to December 7, 2018. Nagoya City citizens discussed five themes about their city's future. We got 15,199 page views, visits from 798 participants, 157 registered participants, and 432 submitted opinions. We established two phases: a 30-day-discussion and a 7-day phase for agreeing to the summarized ideas.

Our main objective is to gather opinions and discussions for a midterm draft of the Nagoya-city Next-Generation Comprehensive Plan, generated by the Nagoya municipal assembly, the local government, and its offices. The plan has five main themes: Theme 1: Human rights and diversity, Theme 2: Secure child care, Theme 3: Disaster prevention, Theme 4: City environment, and Theme 5: Attractiveness for industry and the world.

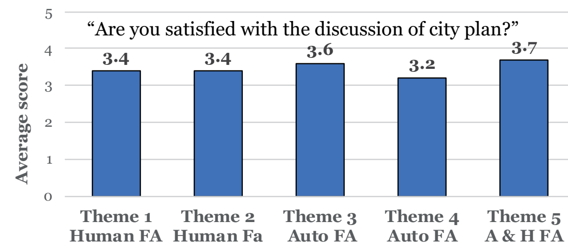
Themes 1 and 2 were facilitated by expert facilitators. Themes 3 and 4 were facilitated only by automated facilitation agents. Theme 5 was facilitated by cooperation between humans and agents.

Figure 4 shows the result for (A) the number of posts for each theme. Figure 5 shows the result for (B) the user satisfaction scores. In (A), the themes facilitated by the automated facilitation agent (Auto FA) obtained more posts from the participants, meaning that the automated facilitation agent (Auto FA) incentivized participants to submit more opinions. In (B), the satisfaction scores are almost the same among all of the themes, suggesting that users can experience satisfying discussions even if they are facilitated by the automated facilitation agent (Auto FA).

Theme	Postings			
	All	Human FA	Auto FA	Participants
1: Human FA	81	43	0	38
2: Human FA	56	21	0	35
3: Auto FA	88	0	24	64
4: Auto FA	70	0	18	52
5: A & H FA	137	17	21	99
Sum	432	81	63	288

Human FA is Human Facilitator. Auto FA is Automated Facilitation Agent numbers indicated that Auto FA successfully extracted more posts than Human

Figure 4: Experiment result: Posting



Questionnaire (N=20) results of "are you satisfied with the discussion of the city plan?". Satisfaction scores on discussion by Auto FA were satisfied at same level as discussion by Human FA.

Figure 5: Experiment result: Satisfaction Score

Figure 6: Actual case of successful automated facilitation agent

Figure 6 shows an actual case where our automated facilitation agent successfully facilitated a discussion among civilians. Issue 1 was raised by the participants. The automated facilitation agent identified this post as an Issue. Then he/she asked "What can we do to solve it?" Then a participant posted idea 1. The automated facilitation agent identified this post as an idea and raised an issue to deepen the idea. Then a participant posted idea 2. The automated facilitation agent works very efficiently. In particular, identifying posting type is accurate because of deep learning technology and well-trained data. This case was successful, but unfortunately, we also experienced some failures.

4 Related Work

We were inspired to enter this area by several ongoing intriguing projects, of which the following are representative. The goal of the Climate CoLab [Malone and Klein, 2007], which is one of the most famous web-based collective intelligence projects, is to harness the collective intelligence of thousands of people worldwide to address global climate change. Like Wikipedia and Linux, MIT CCI developed a crowdsourcing platform where citizens work with experts to create, analyze, and select detailed proposals that tackle climate change. This system defined several steps, including "proposal creation," "finalist selection," "proposal revisions," "voting," and "presentations to potential implementers" to integrate innovative opinions with crystalized ideas that are implementable. Deliveratorium [Iandoli *et al.*, 2007] is another project where people submit ideas by following an argumentation map, which is a kind of discussion structure through which people frame their ideas. With structured argumentation maps, Deliveratorium makes it possible to clearly show the entire relations among ideas and opinions. Such structuring can be done even if the opinions are completely divided.

One of the related fields on extracting discussion structures is argumentation mining. Argumentation mining aims to identify argument structures in natural language texts. For example, many studies in this field extract structures from essays [Stab and Gurevych, 2014b; Nguyen and Litman, 2016][Stab and Gurevych, 2017], reviews [Kim, 2014], and legal texts [Palau and Moens, 2009] in the same way as we extract structures from online discussions. The main difference is that we use extracted stuctures to facilitate discussion in realtime manner in order to make real discussion coherent while these work basically just extract structure from the static documents.

Wong and Aikin[Wong and Aiken, 2003] worked about automated facilitator but their approach is to provide a predefined framework that is a kind of static guide for online discussion, which they call it as "automated facilitator". This is totally different from our facilitation agent. Tavanapour and Bittner[Tavanapour and Bittner,] proposed an automated facilitator but their experiment is based on Wizard-of-Oz method. There is not a real software/program implementation as like ours. Adla [Adla *et al.*, 2011] proposed some toolkits for human facilitator. We also provided such supporting functions for human facilitator in the system. Limayem[Limayem, 2006] proposed an automated falitation framework where they proposed a predefined framework to guide statically online discussion. Compared with these earlier works, our facilitation agent can identify the semantic structure of discussion, analyze it, and then post its own message to the discussion board system. This feautre is more flexible than the above researches in the past.

5 Conclusion

We presented our current implementation of a crowd-scale discussion support system based on an automated facilitation agent, which extracts discussion structures from text discussions, analyzes them, and posts facilitation messages. We did a large-scale experiment with Nagoya City ' s local government in which our automated facilitation agent worked quite well. Much important future work remains. We would like to investigate the extent to which the agent can be biased, which is related to ELSE (Ethics, Law, Social-issue, and Economics). Although our current agent is implemented to manage equitable facilitations, but perhaps agents might be involved in biased facilitation. If so, an interesting topic is investigating how much bias people can accept. Since our agents cannot currently generate issues, they are given beforehand or generated by participants. This is required for constructive discussions so that participants are aware of the different aspects of the discussion theme. Several techniques might achieve this goal.

Acknowledgement

This work was supported by the JST CREST fund (Grant Number: JPMJCR15E1). I would like to express the deepest appreciation to all people who participated in our social experiment inlcluding people in Nagoya-city local goverment and civilian in Nagoya-city, and also people who are related to our research and development.

References

[Adla *et al.*, 2011] A. Adla, P. Zarate, and J.-L. Soubie. A proposal of toolkit for gdss facilitators. *Group Decision and Negotiation*, 20(1):57–77, Jan 2011.

[Iandoli *et al.*, 2007] Luca Iandoli, Mark Klein, and Giuseppe Zollo. Can we exploit collective intelligence for collaborative deliberation? the case of the climate change collaboratorium. 2007.

[Ito and Shiramatsu, 2018] Takayuki Ito and Shun Shiramatsu. Consensus support device and program for consensus support device, 2018. Japan Patent 2018-148665.

[Ito *et al.*, 2014] Takayuki Ito, Yuma Imi, Takanori Ito, and Eizo Hideshima. Collagree: A faciliator-mediated large-scale consensus support system. *Collective Intelligence 2014*, 2014.

[Ito *et al.*, 2015] Takayuki Ito, Yuma Imi, Motoki Sato, Takanori Ito, and Eizo Hideshima. Incentive mechanism for managing large-scale internet-based discussions on collagree. *Collective Intelligence*, 2015, 2015.

[Ito, 2018] Takayuki Ito. Towards agent-based large-scale decision support system: The effect of facilitator. In *The 51st Hawaii International Conference on System Sciences (HICSS2018)*, 2018.

[KAWASE *et al.*, 2018] Satoshi KAWASE, Takayuki ITO, Takanobu OTSUKA, Akihisa SENGOKU, Shun SHIRAMATSU, Tokuro MATSUO, Tetsuya OISHI, Rieko FUJITA, Naoki FUKUTA, and Katsuhide FUJITA. Cyber-physical hybrid environment using a largescale discussion system enhances audiences' participation and satisfaction in the panel discussion. *The IEICE Transactions on Information and Systems*, E101.D(4):847–855, 2018.

[Kim, 2014] Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

[Kunz and Rittel, 1970] Werner Kunz and Horst WJ Rittel. Issues as elements of information systems. Technical report, 1970. CiteSeerX 10.1.1.134.1741.

[Lample *et al.*, 2016] Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, and Chris Dyer. Neural architectures for named entity recognition. In *Proc. of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 260–270. Association for Computational Linguistics, 2016.

[Lawrence and Reed, 2017] John Lawrence and Chris Reed. Mining argumentative structure from natural language text using automatically generated premise-conclusion topic models. In *Proc. of the 4th Workshop on Argument Mining*, pages 39–48, 2017.

[Limayem, 2006] Moez Limayem. Human versus automated facilitation in the gss context. *SIGMIS Database*, 37(2-3):156–166, September 2006.

[Malone and Klein, 2007] Thomas W Malone and Mark Klein. Harnessing collective intelligence to address global climate change. *Innovations: Technology, Governance, Globalization*, 2(3):15–26, 2007.

[Malone, 2018] Thomas W. Malone. *Superminds: The Surprising Power of People and Computers Thinking Together*. Little, Brown and Company, 2018.

[Nguyen and Litman, 2016] Huy Nguyen and Diane Litman. Context-aware argumentative relation mining. In *Proc. of the 54th Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1127–1137, 2016.

[Nishida *et al.*, 2017] Tomohiro Nishida, Takayuki Ito, Takanori Ito, Eizo Hideshima, Shunpei Fukamachi, Akihisa Sengoku, and Yumika Sugiyama. Core time mechanism for managing large-scale internet-based discussions on collagree. In *the Proc. of the 2nd IEEE International Conference on Agents (IEEE ICA2017)*, 2017.

[Nishida *et al.*, 2018] Tomohiro Nishida, Takanori Ito, and Takayuki Ito. Verification of effects using consensus-building support system in continuous workshops for city development. *Journal of the Science of Design*, 11 2018.

[Palau and Moens, 2009] Raquel Mochales Palau and Marie-Francine Moens. Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 98–107. ACM, 2009.

[Sengoku *et al.*, 2016] Akihisa Sengoku, Takayuki Ito, Kazumasa Takahashi, Shun Shiramatsu, Takanori Ito, Eizo Hideshima, and Katsuhide Fujita. Discussion tree for managing large-scale internet-based discussions. *Collective Intelligence*, 2016, 2016.

[Stab and Gurevych, 2014a] Christian Stab and Iryna Gurevych. Annotating argument components and relations in persuasive essays. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics*, pages 1501–1510, 2014.

[Stab and Gurevych, 2014b] Christian Stab and Iryna Gurevych. Identifying argumentative discourse structures in persuasive essays. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 46–56, 2014.

[Stab and Gurevych, 2017] Christian Stab and Iryna Gurevych. Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659, 2017.

[Tavanapour and Bittner,] Navid Tavanapour and Eva Bittner. Automated facilitation for idea platforms: Design and evaluation of a chatbot prototype. In *39th International Conference on Information Systems (ICIS) 2018*, oct.

[Wong and Aiken, 2003] Zachary Wong and Milam Aiken. Automated facilitation of electronic meetings. *Information and Management*, 41:125–134, Dec 2003.

[Yamaguchi *et al.*, 2018] Naoko Yamaguchi, Takayuki Ito, and Tomohiro Nishida. A method for online discussion design and discussion data analysis. In *The 13th International Conference on Knowledge, Information and Creativity Support Systems (KICSS-2018)*, 2018.