

ISSN 2758-075X

東京国際工科専門職大学紀要

第3巻 第1号

**Bulletin of International Professional University of
Technology in Tokyo**

2024年 3月

東京国際工科専門職大学

目 次

論文

- 帰国子女 TOEFL iBT 学習者を対象とした認知バイアス効果が
学業成績に与える影響についての検証
菅谷 孝義…………… 1

研究ノート

- ChatGPT のデータ解析系授業への影響について
三宅 茂樹…………… 15

STEAM 教育に適した参加度向上のための環境開発

- 兒玉 賢史・爰川 知宏・武本 充治・辻野 雅之・
藤井 竜也・山本 裕…………… 22

「ヒト・モノ協調型のスマートモビリティ社会」に関する研究報告

- 辻野 雅之・内埜 大斗・水上 憲明・山口 直彦…………… 29

Sound AR サービス「Locatone (ロケトーン)」を用いた企業連携実習の現状と展望

- 小野 憲史・水上 恵太・安彦 剛志・青山 龍・八木 泉…………… 40

留学生の就職成功に求められる日本語能力・人材育成の効果的アプローチ

- 専門学校及び日本語学校での取り組み—
山口直彦・伊勢みゆき…………… 49

生成 AI を活用した未知の領域の学習への試み

- 量子コンピューティングを例にして—
廣峯 優希・富山 哲男・宮崎 淳…………… 66

調査研究

- LLM の評価手法の提案と学習支援効果
安藤 優希・富山 哲男・山下 一郎…………… 82

- 『東京国際工科専門職大学紀要』規程…………… 97

- 『東京国際工科専門職大学紀要』執筆要項…………… 99

Contents

Articles

- The effects of cognitive bias on academic success: The case for returnee learners of TOEFL iBT®
----- Takayoshi Sugaya --- 1

Research Notes

- Impact of ChatGPT on Data Analysis Courses
----- Shigeki Miyake --- 15

- Development of an environment suitable for STEAM education to improve participation
----- Satoshi Kodama, Tomohiro Kokogawa, Michiharu Takemoto,
Masayuki Tsujino, Tatsuya Fujii and Hiroshi Yamamoto --- 22

- Research Report on “Smart Mobility Society based on Human-Object Coordination”
----- Masayuki Tsujino, Hiroto Uchino, Noriaki Mizukami,
and Naohiko Yamaguchi --- 29

- A prospect for corporate collaborative training using the Sound AR service “Locatone”
----- Kenji Ono, Keita Mizukami, Takeshi Abiko, Ryu Aoyama,
and Izumi Yagi --- 40

- Effective Approaches to Human Resource Development and Japanese Language
Proficiency Required for Successful Employment of International Students
----- Naohiko Yamaguchi and Miyuki Ise --- 49

- Attempts to learn the unknown by using generative AI: For example, Quantum Computing
----- Yuki Hiromine, Tetsuo Tomiyama, and Jun Miyazaki --- 66

Research Reports

- Proposal for LLM Evaluation Methods and Learning Support Effectiveness
----- Yuki Ando, Tetsuo Tomiyama and Ichiro Yamashita --- 82

【論 文】

The effects of *cognitive bias* on academic success: The case for returnee learners of TOEFL iBT®

Takayoshi Sugaya

Abstract : The goal of this study is primarily to find out ways to help students learn to recognize that various *cognitive biases* could impact their thinking and make adjustments as needed. Specifically, the study aims to identify significant biases and search *debiasing* methods that could promote the acceptance of returnee learners of English who have studied overseas. Using the psychological approach, the author investigated eleven *cognitive bias* traits assumingly characteristic of returnee learners of TOEFL iBT® and correlations between *cognitive bias* factors and their academic success. Multiple regression analysis was used to examine detrimental effects of *cognitive bias* on academic performance. Two main findings were obtained for correlations between 11 *cognitive biases* and participants' test scores. *Choice-supportive bias* and *outcome bias* had positive correlations with learners' test scores in terms of learner psychology while *reactance bias* was found to have a significantly negative impact on learning effects. It was found out that *cognitive biases* could work both positively and negatively and that certain *cognitive biases* and academic success are compatible. These findings have some pedagogical implications. Teachers can make use of the findings of the study in helping learners enhance learners' motivation, confidence and self-esteem by positively exploiting certain *cognitive biases* in their instructions and mitigate detrimental bias traits which could adversely influence learners' academic performances with careful *debiasing* guidance.

Keywords : cognitive bias, debias, TOEFL iBT®, returnee

1. Introduction

In terms of learner autonomy, *cognitive bias* is the opposite of *metacognition*. It is possible that learners are making cognitive errors in their own thinking and it is hindering their learning. *Metacognition* works best when learners incorporate awareness of possible bias and errors in their cognitions. In other words, in order to be successful in learning, it is not enough to just be metacognitive, but learners must also be aware of possibilities they might be relying on a biased interpretation of their learning. On the premises that they have negative influences on their learning effects in terms of learner psychology, this paper empirically examines *cognitive bias* traits relevant with academic success in the context of TOEFL iBT® and explores *debiasing* methods that could promote the acceptance of returnee learners and ways to guide them to academic success on the basis of the previous literature in cognitive psychology.

1.1 Cognitive bias

First of all, it is necessary to clarify and elaborate on the concept of “*cognitive bias*” for vocabulary knowledge. By definition, a *cognitive bias* is a systematic pattern of deviation from norm or rationality in judgment (Haselton et al., 2005). First coined by Amos Tversky and Daniel Kahneman (1974), cognitive biases are considered to be errors in thinking that can lead to a misinterpretation of information, affecting the accuracy of our decisions and the rationality of our judgements. Since the introduction of the concept of *cognitive bias*, there has been a considerable amount of research conducted into this psychological field. Various researchers and psychologists have weighed on what they believe constitutes a *cognitive bias*. There are over 50 *cognitive biases* that psychologists consider when they examine thinking (See Table 1).

Table 1 Taxonomy of cognitive biases in psychology

Decision-making and behavior
<i>Bandwagon effect / Confirmation bias / Congruence bias / Contrast effect / Distinction bias / Endowment effect / Extreme aversion / Focusing effect / Framing effect / Illusion of Control / Impact bias / Information bias / Loss aversion / Mere exposure effect / Neglect of probability / Omission bias / Outcome bias / Ostrich effect / Planning fallacy / Reactance / Status quo bias / Zero-risk bias</i>
Probability
<i>Ambiguity effect / Anchoring / Attentional bias / Availability heuristic / Availability cascade / Clustering illusion / Conjunction fallacy / Exaggerated expectation / Gambler's fallacy / Hawthorne effect / Hindsight bias / Hot-hand fallacy / Illusory correlation / Ludic fallacy / Neglect of prior base rates effect / Ludic fallacy / Neglect of prior base rates effect / Optimism bias / Overconfidence effect / Positive outcome bias / Primacy effect / Recency effect / Repetition bias / Stereotyping</i>
Attribution
<i>Actor-observer bias / Dunning-Kruger effect / Egocentric bias / Empathy gap / Forer effect / Fundamental attribution error / Halo effect / Illusion of asymmetric insight / Ingroup bias / Projection bias / Restraint bias / Self-serving bias / Self-fulfilling prophecy / Trait ascription bias</i>
Memory
<i>Beneffectance / Choice-supportive bias / Continued influence effect / Consistency bias / Cryptomnesia / Egocentric bias / Suggestibility</i>

Note: On the basis of the research book by Kisak (2015), cognitive bias items were categorized into four subscales by the author.

Among over 50 cognitive bias theories, 11 bias theories were selected by the author in relevance with participants. The current study merely focuses on eleven biases in bold which the author assume are most commonly associated with returnee learners' psychology from the pedagogical perspective (See Table 2). On speculation, those eleven selected cognitive bias traits seem to be most typical of returnee learners of English in terms of learner psychology, namely, overconfidence, pridefulness, self-reliance, freedom of choice, wishful thinking, causal attribution and memory errors, and consequently seem to have an impact on their academic performances.

Table 2 *Leading proponents & Definitions of 11 cognitive biases chosen*

<i>Decision-making & behavior</i>		
1. <i>Outcome bias</i>	Baron, J., & Hershey, J. C.	The decision based on the outcome of previous events, without regard to how the past events developed (Baron & Hershey, 1988).
2. <i>Ostrich effect</i>	Galai, D. & Sade, O.	The avoidance of apparently risky financial situations by pretending they do not exist (Galai & Sade, 2006).
3. <i>Planning Fallacy</i>	Kahneman, D., & Lovallo, D.	The tendency to underestimate the time, costs, and risks of future actions and at the same time overestimate the benefits of the same actions. (Kahneman & Lovallo, 2003)
4. <i>Reactance</i>	Brehm, J. W.	The motivation to regain a freedom after it has been lost or threatened (Brehm, 1989).
5. <i>Status quo bias</i>	Samuelson, W., & Zeckhauser, R.	The preference for the maintenance of one's current or previous state of affairs (Samuelson & Zeckhauser, 1988).
Probability		
6. <i>Exaggerated expectation</i>	Hilbert, M.	The tendency to expect or predict more extreme outcomes than those outcomes that actually happen (Hilbert, 2012).
7. <i>Optimism bias</i>	Weinstein, N. D.	The tendency for people to believe that they are less likely to experience negative events and more likely to experience positive events than are other people. (Weinstein, 1980)
Attribution		
8. <i>Dunning-Kruger effect</i>	Dunning, D., Johnson, K., Ehrlinger, J., and Kruger, J.	A cognitive bias in which people wrongly overestimate their knowledge or ability in a specific area (Dunning et al., 2003).
9. <i>Self-serving bias</i>	Larson, J. R.	The tendency to perceive oneself in an overly favorable manner (Larson, 1977).
Memory		
10. <i>Choice supportive bias</i>	Lind, M., Visentini, M., Mäntylä, T., & Del Missier, F.	The tendency to retroactively ascribe positive attributes to an option one has selected and/or to demote the forgone options (Lind et al., 2017).
11. <i>Continued influence</i>	Johnson, H. M., & Seifert, C. M.	The tendency for misinformation to continue to influence memory and reasoning about an event (Johnson & Seifert, 1994).

1.2 Debiasing Approach

“*Debiasing*” is also referred to using other terms, such as *Cognitive Bias Modification*, *cognitive bias mitigation*, *cognitive bias reduction*, is the prevention and reduction of the negative effects of cognitive biases, particularly with respect to judgment and decision making (Morewedge et al., 2015). To put it differently, *debiasing* is a process through which the influence of *cognitive*

biases is reduced, generally with the goal of helping people think in a more rational and optimal manner. Through a debiasing process, a person may learn or adopt better strategies by which to make judgments and decisions (Larrickcs, 2004). There are three general approaches to debiasing judgment and decision making, and the costly errors with which biased judgment and decision making is associated: changing incentives, nudging, and training. Since part of the research aim is to identify significant biases and search effective *debiasing* methods from the educational point of view, the discussion involves prescriptive *debiasing* guidance referring to the debiasing theory as proposals in academic disciplines. Participants were asked to examine their own biases in the questionnaire before conducting lessons in the hope that this process would help them develop awareness of their own biases in their thinking and in turn help guide their future thinking and behaviors.

2. Method

2.1 Research participants

High school participants were drawn from among those the author individually gave instructions on TOEFL iBT[®] through a number of private lessons (12~36 lessons). The participants were 33 high school Japanese students (16 males, 17 females), between the age of 15 and 18, who attended either local high school (no ESL or EFL classes involved) or international school overseas attended or unattended by parents, where all class lessons were conducted in English. They all attended local high schools or international schools in such countries as U.S.A. (17), Thailand (5), Germany (4), Canada (3), U.K (1), Brazil (1), France (1), Singapore (1). The students completed a self-report questionnaire which focused on finding out the degree of cognitive bias traits in learning. They all needed TOEFL iBT[®] score in order to prove their competence in English in applying for Japanese universities as returnee applicants. The number of TOEFL iBT[®] they actually took vary ranging from three to six times.

2.2 Academic achievements and Data Collection

As a measure to assess participant learners' leaning effects and academic success, the TOEFL iBT[®] test was employed for the following reasons. First, they all took the test multiple times and the test was suited to know their learning effects. Next, TOEFL iBT[®] was most commonly taught by the author in instructing returnee high school learners and their data was the largest in number. Besides, TOEFL iBT[®] is a high-stakes test which measures the abilities of test-takers in four language skills and suited to know their comprehensive English language skills at the higher education level. As for academic achievements, all participants were assessed individually according to their test results so that their learning effects can be measured. Score data were drawn from among those who took the test more than twice and their learning period varied from three months to twelve months. The experiment took place at a preparatory school where the author taught as a part-time English instructor. A series of private lessons (12~36) were given either face to face or online (Skype)

individually. Participants were asked to report their latest score detailed in each section of TOEFL iBT® composed of four sections in which 30 is given as a full score: reading, listening, speaking and writing, totaling 120 as a full score. The author kept track of every participant's test results in detail and made detailed analysis of their academic improvements in scores as learning effects. The highest total scores of each participant were used for the multiple regression analysis as measures of their academic success. Their total scores of TOEFL iBT® ranged from 60 to 111. The mean score was 82.1.

2.3 Ethics Statement

The aim of this questionnaire survey is to measure learners' inclination and their learning effects. The questionnaire is answered on online questionnaire form (Google Form). The author asked each respondent on the first page of the form a consent to the use of survey results and their score data for the purpose of research only. It was explained to them that careful considerations were taken in order to protect respondents' privacy and survey results were to be anonymous in the paper. All participants gave their informed consents before taking part in the study.

2.4 Measures for Learner's Cognitive Bias

After providing consents, the participating learners were asked to complete the questionnaire and answer questions provided by the author which specifically focused on finding out the degree of cognitive bias in learning (See Table 3). The survey questionnaire is generally based on the work of P. F. Kisak (2015). In order to measure the degree of cognitive bias, eleven items, which the author assume are relevant in autonomous learning, were used in the questionnaire. The scales of the questionnaire comprised the five-point Likert type statements ranging from [1], the least accurate, to [5] the most accurate. The goal of these questions was to help students think about ways their own cognitive errors may be contributing to their learning and scores. Cognitive bias question items in the questionnaire were reduced to about one-fifth in number because fewer number of questions would not adversely affect participants' motivation in learning and also answering over 50 cognitive bias item questions might have affected the qualities of their answers.

Table 3 11 Questionnaire item

Decision-making and behavioral biases
“I place much importance on the final outcome rather than an assessment of the time and quality of previous learning.” (<i>Outcome bias</i>)
“I tend to avoid or ignore unfavorable information including negative feedbacks and undesirable outcomes.” (<i>Ostrich effect</i>)
“I tend to underestimate the time needed to complete future tasks in learning.” (<i>Planning fallacy</i>)
“When negative feedbacks are given or learning methods are forcefully regulated by the teacher, I tend to react against suggestions given even if they seem to be good ones and try to maintain my own style of learning.” (<i>Reactance</i>)
“I prefer to maintain the current baseline and try not to make a change in learning even when better alternatives are available.” (<i>Status quo bias</i>)
Biases in probability and belief
“I tend to expect more extreme outcomes than those outcomes that actually happen.” (<i>Exaggerated expectation</i>)
“I tend to believe that I am less at risk of experiencing a negative outcome and estimate its chances lower than it might actually happen.” (<i>Optimism bias</i>)
Attributional biases
“I often overestimate my ability in learning.” (<i>Dunning-Kruger effect</i>)
“I feel irresponsible for undesirable outcomes.” (<i>Self-serving bias</i>)
Memory errors
“I attribute positive outcomes to options I chose.” (<i>Choice supportive bias</i>)
“I tend to continue to pursue learning methods initially taken even when they later proved to be false or obsolete.” (<i>Continued influence</i>)
<i>Note:</i> The actual questionnaire was rearranged in random order without subscales or cognitive bias types in parentheses.

3. Results

Multiple linear regression was used to identify cognitive bias factors from 11 predictor variables which significantly affected participants' own scores in TOEFL iBT® (dependent variable). When processing data, the author used the statistical software BellCurve for Excel. Using the Stepwise method, model that includes the variables that are important in the prediction and excludes the ones that have only a trivial effect was found (See Table 4 for a summary of the results).

Table 4 Descriptive Statistics of Multiple Regression

<i>r</i>	<i>r</i> ²	<i>r</i> ² Adjusted	<i>n</i>			
.66	.58	.44	33			
<i>ANOVA</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Sig. (p)</i> ^a	
Regression	5	3387.7	677.5	4.2	.006**	
Residual	27	4367.9	161.8			
Total	32					
<i>Coefficients</i>						
Variables	β	<i>SE</i>	<i>t</i>	<i>p</i>	Lower 95%	Upper 95%
<i>Reactance</i>	-8.2	2.7	-3	.006**	-13.8	-2.6
<i>Outcome bias</i>	7.8	2.3	3.3	.003**	3	12.6
<i>Choice-supportive bias</i>	6.1	2.5	2.4	.02*	0.9	11.2
<i>Optimism bias</i>	3.3	2	1.6	.11	-0.8	7.3
<i>Exaggerated expectation</i>	3.5	2.1	1.6	.12	-0.9	7.8
Constant	31.9	14.1	2.3	.03*	3.1	60.8

Note : ^a Dependent Variable: Total Score, *: $p < .05$, **: $p < .01$,

The findings obtained for the correlation between 11 cognitive biases and participants' test scores were satisfactory (See Table 3). The overall regression was statistically significant, $F(5, 27) = 4.2$, $r^2 = .58$, $p = .006$. It was found out that, out of all examined variables, three variables were found to be significant; Variable *Choice-supportive bias* ($p = .02$) was fairly significant, and variables *Reactance* ($p = .006$) and *Outcome bias* ($p = .003$) were highly significant which most likely contribute to the overall variability of academic success, whereas two remaining predictor variables: *Optimism bias* ($p = .11$), *Exaggerated expectation* ($p = .12$) did not significantly predict high score. Namely, two variables *Outcome bias* ($\beta = 7.8$, $t = 3.3$) and *Choice-supportive bias* ($\beta = 6.1$, $t = 2.4$) had positive correlations with learners' test scores, on the other hand, one variable, *Reactance bias* ($\beta = -8.2$, $t = -3$) had a significantly negative influence on their scores.

4. Discussion

4.1 Choice-supportive bias

The results revealed that certain *cognitive biases* could work positively and negatively. I should then like to move on to discuss the ambivalent results of significant variables and interpret their positivity and negativity. The question "I attribute positive outcomes to options I chose." was answered before they knew the test result. They did not make any attributions yet to their TOEFL iBT® scores at the time. Accordingly, each participant's self-analysis at the time could be construed as their innate or general tendency. *Choice-supportive bias* had a rather positive influence on participants' test results ($p = .02$). Contrary to the author's prediction, the result showed that the more they attribute positive outcome to options they chose, the higher scores they obtained.

Concerning this positivity of the bias shown, the author speculates that this bias positivity can be construed in accordance with the *positive illusion theory* as possible explanations. Greenwald (1980) proposes that people also hold positive illusions because such beliefs often enhance their productivity and persistence with tasks on which they might otherwise give up. Boiney et al. (1997) also propose that “motivated decision makers bias their judgments more or less as needed to support the desired conclusion and motivated decision makers exhibit confidence bolstering and thereby remain at least as confident as non-motivated decision makers in their biased estimates” (1997, p.1). Regarding this positivity of cognitive biases, Cummins & Nistico (2002) claim that “well-being homeostasis is controlled by positive cognitive biases pertaining to the self. Most particular in this regard are the positive biases in relation to self-esteem, control and optimism” (2002, p.37). Concerning positivity of *choice-supportive bias* shown in the current research, it is possible that memories of chosen alternatives affected participant learners’ sense of well-being. Positive illusions are a form of self-deception or self-enhancement that feel good, maintain self-esteem, or avoid discomfort, at least in the short term. Positive illusions have been commonly understood as one of the apparent effects of self-enhancement, a desire to maximize the positivity of one’s self-views and “self-enhancing biases often make people feel good about themselves and have other short-term benefits” (Leary, 2007, p.324). In a positive light, *positive illusions* can be important contributors to one’s positive feelings and can make a good outcome in a short-term learning. Accordingly, in the case of participant learners who worked on the test relatively in a short-term, positivity of *choice-supportive bias* shown in relation to their academic performances can be attributed to short-term effects of *positive illusion theory* in terms of self-esteem, optimism and self-enhancement as suggested in the previous literature.

4.2 Outcome bias

Outcome bias also significantly had a positive influence on test results ($p = .003$). The question “I place much importance on the final outcome rather than an assessment of the time and quality of previous learning.” was answered before conducting lessons. They did not make any attributions yet to their TOEFL iBT[®] in the questionnaire. The means that each participant’s self-analysis scales concerning this bias question are also rather their innate or general trait. The result found that the more participants put importance on result, the higher their test results got. They put more stress on test results and processes were less significantly considered in learning. As described in Table 2, *Outcome Bias* is the tendency to evaluate decisions solely based on the outcomes. A decision is based on the outcome of previous events, without regard to how the past events developed. and “people take outcomes into account in a way that is irrelevant to the true quality” (Baron & Hershey, 1988, p.570). However, the *outcome bias* trait, which is thus commonly categorized as a negative trait, worked positively in this empirical study. Considering the fact that participants were returnee learners who were applying for universities with TOEFL scores, there is no doubt that they were desperately in need of good test results. This might explain why they valued results more than learning processes and more goal-oriented rather than process-oriented. Hence, it can be surmised that in the case of a relatively short-term learning, it is possible that *outcome bias* mindset worked

positively as *choice-supportive bias* did.

4.3 Reactance Bias

The result of multiple regression test showed that *reactance bias* was the most influential variable negatively influencing participants' TOEFL iBT® scores. Degrees of *reactance bias* among participant learners were measured based on their answer scales to the question: "When negative feedbacks are given or learning methods are forcefully regulated by the teacher, I tend to react against suggestions given even if they seem to be good ones and try to maintain my own style of learning." The result indicated that the more negative feedbacks they receive and forcefully regulated, the more reactant and persistent they become in learning. In addition to this, the more reactant they became, the lower their scores became. This means that negativity of feedbacks and forcefulness on learning styles obviously had adverse effects on their academic performances.

Reactance is an unpleasant motivational arousal that emerges when people experience a threat to or loss of their free behaviors (Brehm, 1989). The concept of psychological *reactance* was originally formulated by psychologist Jack Brehm. He explains that *reactance* occurs when "people are motivated to restore specific behavioral freedoms that are threatened or taken away from them" and "people become motivationally aroused by a threat to or elimination of a behavioral freedom (1989, p.72). In short, reactance can be defined as the motivation to regain a freedom after it has been lost or threatened. It causes individuals to rebel against the pressure they are placed under. Persuasion, enforcement, prohibition and instruction all can examples of threats to the freedom to act as desired, and this is where reactance comes into action. In terms of learner psychology, *reactance* is a response of noncompliance and serves as a motivator to restore one's freedom. Students do not want to be told that they have been studying in the wrong way. In such cases, giving negative feedbacks on their competence in learning and regulating their learning styles in a forceful or compulsive manner, regardless of its accuracy, can be risky in the sense that getting straight to the point might incur disagreements and conflicting situations might lead to some adverse outcomes for both sides.

4.4 Prescriptive Debiasing

Taking potential risks into consideration, teachers must find ways to reduce this *reactance* trait with effective debiasing techniques because it is difficult for individuals to recognize and correct biases (Kahneman, 2003). *Debiasing* is the art of reducing biases in human thinking, by finding a variety of useful bias-reducing techniques and most importantly, "*debiasing* requires intervention" (Larrick, 2004, p.318). In education, however, interventions must be interpreted positively by the learning side. Overall findings of the previous research show that cognitive *debiasing* works in some cases and can be helpful when it comes to a large range of *cognitive biases* (Lilienfeld et al., 2009). The use of appropriate training, interventions, and *debiasing* techniques can reduce some cognitive biases, to some degree, in some situations. However, in order to reduce *cognitive biases* debias successfully in the individual, we should engage in a proper *debiasing* process because *debiasing* is not always an easy process and finding the appropriate *debiasing* approach to use

in a certain situation can be a difficult process because a *debiasing* approach that works well in one situation might fail in another (Sherbino et al., 2014). In the school setting, it is particularly important for the teaching side to be able to engage in proper *debiasing*. We should keep it in mind that the effectiveness of debiasing varies based on a large range of factors including the person that we are trying to debias and the specific bias that we are trying to reduce. There are three approaches considered to be effective to *debiasing* cognitive errors: to develop awareness of the bias, to challenge old beliefs using different perspectives and to elicit external feedback (Soll et al., 2015). Each approach has strengths and weaknesses and there is a limitation in using domain-general approaches as general-purpose prescription. In addition to general *debiasing* approach, there are also some more specialized *debiasing* techniques that are meant to be used on a narrower range of biases. Such techniques can sometimes be more effective than generalized *debiasing* strategies. Once specific bias trait is identified, domain-specific treatments or prescription will be more effective. In light of this, I should like to explore possibilities of prescriptive methods, to mitigate *reactance* by learners referring to two *debiasing* theories: namely, *nudge*, and *reverse psychology*.

A “*nudge*” is an intervention in the decisional context that steers people’s decisions, actions, and behaviors by acting on their cognitive biases. Overall, *nudges* can selectively work for some people, in some domains, some time (Thaler & Sunstein, 2009). The *nudge* theory is the idea of altering people’s behavior in a predictable way without forbidding any options. Instead of choosing an option, learners can be assigned by instructors. Providing learners with assigned options can be beneficial for some reasons. First, since *nudges* are not mandatory, students should feel that the nudge respects their ability to choose. Specifically, we should make sure that the people who are affected by the *nudge* feel that it doesn’t eliminate their autonomy and right to choose freely. When learners feel like they have a choice it can increase their sense of freedom and reduce *reactance*. Second, making a choice or having a choice made for you by other people in your best interest can prompt memory attributions that support that choice (Mather et al., 2003). Third, providing learners with outcomes and details of what will happen if they follow the assigned option will create a sense of comfort and give learners a sense of control over the situation.

What is more, *Reactance* can occur when an individual senses that someone is trying to compel them to do something and often, the individual will offer resistance and attempt to extricate themselves from the situation. It is a personality characteristic and some individuals are naturally high in *reactance*. In such cases, *reverse psychology* is also a technique involving the assertion of a belief or behavior that is opposite to the one desired, with the expectation that this approach will encourage the subject of the persuasion to do what is actually desired. The one being manipulated is usually unaware of what is really going on. *Reverse psychology* relies on the psychological phenomenon of *reactance*, in which a person has a negative emotional reaction to being persuaded, and thus chooses the option which is being advocated against (MacDonald et al., 2011). In teaching, *Reverse psychology* can be implemented from two perspectives. On the one hand, it can be used as a manipulative persuasion tactic in a negative manner. Alternatively, it can also be used as a helpful method to benefit relationships.

4.5 Educational Implications

The findings obtained in the current study have some educational implications. Regarding the positive use of cognitive bias, for the purpose of enhancing learners' motivation, confidence, self-esteem, *choice-supportive bias* can be positively exploited to affect learner's decisions and can be used to the teacher's advantage. There might be some cases where someone needs to tell learners that they are not being objective and their tendency to cling to his or her decision too long is damaging the progress. However, when learners stick to a strategy of their choice, it is possible that they have tendencies to tie their self-esteem to their decision-making. The result indicates that participant learners tend to justify their decisions and choices as better over others and still show better results. Hence, rather than explaining the importance and benefits of the process, teachers could intentionally focus on the outcome and how this will affect them. In terms of bias utilities, in a specific case or depending on who we are trying to influence, *outcome bias* can also be used to great effect. In some cases, "optimism can lead people to experience their situation more positively, and overconfidence may help them achieve even unrealistic goals" (Dunning, 2011, p.p.288-9). The research by Robins & Beer (2001) provided evidence that people who have positive illusions may have both short-term benefits and long-term costs. In the current research, since participant learners had to achieve academic success in TOEFL iBT® in a relatively short-term, it is possible that positive illusions or optimism had short-term benefits. Overall, as an educational significance, by understanding the cognitive biases learners have, the teacher could simply justify learners' optimism and choice for them in order to enhance their motivation and self-esteem, and increase the chances of influencing them for the better.

Regarding the negativity of *reactance bias* found and possible debiasing techniques, based on a clear understanding of a learner's current beliefs, and context or domain to be impacted, personalized *nudges* would lead learners to successful behavior change and also to optimal success. For the purpose of mitigating resistance from learners who are high in *reactance* without risks, employing *nudges* would be an effective debiasing technique so that interventions can be interpreted positively. Utilizing *nudges* would also become an opportunity for educators to truly analyze the problem and define what it might take to successfully change individual behavior. Besides, teachers could also employ *reverse psychology* in a bid to exploit *reactance* for their benefit, in an attempt to influence learners to choose the opposite of what is being requested. In order to prevent *reactance* from occurring, employing and applying *reverse psychology* theory in teaching and guidance could be an effective tool in debiasing *reactance* traits, especially for those teachers dealing with returnee learners in adolescence assumingly high in *reactance* who have negative emotional reactions to their advice and tend to persist in their own way of learning. This psychological approach has proven to be particularly effective with adolescents as many of these are prone to rebellious tendencies (MacDonald et al., 2011). Overall, psychological *reactance* occurs as a response to a perceived restriction on our personal freedom and being told not to do something can cause learners to rebel against the situation. The empirical research by Sherbino et al. (2014) provides the evidence that the educational interventions employing cognitive forcing

strategies failed to show any reduction in diagnostic errors and proved to be ineffective in some cases. Therefore, it is necessary for instructors to engage in proper *debiasing* depending on who we are going to debias and be aware of pros and cons of giving succinct feedbacks taking their psychology into consideration and try to curb risks of their emotional rejections (Cohen et al., 2000). In summary, the teaching side should pay more careful attention when dealing with those learners high in *reactance*. Cognitive biases can be directed in a way that improves efficiency and effects in learning as long as teachers ensure that students do not feel controlled, but instead feel empowered to achieve more by accommodating learner autonomy and using healthy constraints.

5. Conclusion

Using the psychological approach, the author investigated eleven *cognitive bias* traits assumingly characteristic of returnee learners and their effects on academic performances. The result found out that *cognitive biases* could work both positively and negatively and that certain *cognitive biases* and academic success are compatible. On the positive side, it is possible that exploiting and utilizing some *cognitive biases* can positively influence learners in terms of motivation, confidence and self-esteem. On the negative side, certain *cognitive biases* could adversely influence academic performance. In such cases, mitigation of detrimental bias traits or debiasing guidance would be effective in teaching. However, when *debiasing*, the teaching side should consider each learner's personal trait and context to maximize its effects without impairing their motivation, confidence and self-esteem. In conclusion, to induce awareness about the need to change and to elicit motivation to change, an ideal approach we should take would be guiding learners with affirmation and encouragement or guiding implicitly and shifting learners for the better.

I hope that this research work provides a helpful framework for those doing coaching and training learners in adolescence, particularly returnee learners. Furthermore, I would be glad if this is of any help to learners in developing awareness of their own biases in their thinking and in turn help guide their future thinking and behaviors.

6. Limitations

In this study, significant cognitive bias factors were tested with eleven biases and effective *debiasing* techniques were discussed, but further researches of effective *debiasing* techniques are needed. There are more cognitive biases types to be considered and more significant cognitive bias traits might be found concerning learner psychology.

Acknowledgments

I would like to express my appreciation to SUNDAI Returnee Club for their continuous support and for their arranging me a number of TOEFL iBT® lessons throughout my research study. Without their constant support, this research study wouldn't have been completed.

References

- [1] Boiney, L. G., Kennedy, J., & Nye, P. (1997). Instrumental bias in motivated reasoning: More when more is needed. *Organizational Behavior and Human Decision Processes*, 72(1), 1-24. <https://doi.org/10.1006/obhd.1997.2729>
- [2] Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of personality and social psychology*, 54(4), 569-579. <https://www.sas.upenn.edu/~baron/papers/outcomebias.pdf>
- [3] Brehm, J. W. (1989). "Psychological Reactance: Theory and Applications", in NA - *Advances in Consumer Research Volume 16*, eds. Thomas K. Srull, Provo, UT: Association for Consumer Research, Pages: 72-75. <https://www.acrwebsite.org/volumes/6883/volum>
- [4] Cohen, G. L., Aronson, J.M., & Steel, C.M. (2000) When Beliefs Yield to Evidence: Reducing Biased Evaluation by Affirming the Self, *Personality and Social Psychology Bulletin* 26 (9),1151-1164. <https://doi:10.1177/01461672002611011>
- [5] Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive cognitive bias. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being*, 3(1), 37–69. <https://doi.org/10.1023/A:1015678915305>
- [6] Dunning, D., Johnson, K., Ehrlinger, J., and Kruger, J. (2003). Why people fail to recognize their incompetence. *Curr. Dir. Psychol. Sci.* 12, 83-87. Doi: 10.1111/1467-8721.01235.
- [7] Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one’s own ignorance. In *Advances in experimental social psychology Vol. 44*, pp. 247-296. <https://rcgd.isr.umich.edu/wp-content/uploads/2018/09/DKE11advance.pdf>
- [8] Galai, D., & Sade, O. (2006). The “ostrich effect” and the relationship between the liquidity and the yields of financial assets. *The Journal of Business*, 79(5), 2741-2759. <https://www.jstor.org/stable/10.1086/505250>
- [9] Greenwald, A.G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*. 35 (7), 603–618. doi:10.1037/0003-066X.35.7.603
- [10] Haselton, M. G., Nettle, D & Andrews, P. W. (2005). *The evolution of Cognitive bias*. In D. M. Buss (Ed.), *The hand book of Evolutionary Psychology*, John wisely & Sons Inc. 724-726. https://www.researchgate.net/publication/308584925_The_Evolution_of_Cognitive_Bias
- [11] Hilbert, M. (2012). Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological bulletin*, 138(2), 211–237. <https://doi.org/10.1037/a0025940>
- [12] Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of experimental psychology: Learning, memory, and cognition*, 20(6), 1420. <https://psycnet.apa.org/doi/10.1037/0278-7393.20.6.1420>
- [13] Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality, *American Psychologist*, 58, 697–720. <https://psycnet.apa.org/doi/10.1037/0003-066X.58.9.697>
- [14] Kisak, P, F. (Ed) (2015). *Cognitive Science & Bias: The Dynamics of Subjective Truth & Human Nature*. CreateSpace Independent Publishing Platform. ISBN-101517533783:
- [15] Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making*. p316–337. Blackwell Publishing. <https://doi.org/10.1002/9780470752937.ch16>
- [16] Larson, J. R. (1977). Evidence for a self-serving bias in the attribution of causality. *Journal of Personality*. 45(3), 430–441. <https://psycnet.apa.org/doi/10.1111/j.1467-6494.1977.tb00162.x>
- [17] Leary, M.R. (2007). “Motivational and emotional aspects of the self”. *Annual Review of Psychology* 58 317 –344. doi:10.1146/annurev.psych.58.110405.085658. PMID 16953794
- [18] Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological re-

- search on correcting cognitive errors promote human welfare? *Perspectives on psychological science*, 4(4), 390-398. [https://www3.nd.edu/~ghaeffel/Lilienfeld2009%20Perspectives%20on%20 Psychological%20Science.pdf](https://www3.nd.edu/~ghaeffel/Lilienfeld2009%20Perspectives%20on%20Psychological%20Science.pdf)
- [19] Lind, M., Visentini, M., Mäntylä, T., & Del Missier, F. (2017). Choice-supportive misremembering: A new taxonomy and review. *Frontiers in psychology*, 8, 2062. p.1-16. <https://doi.org/10.3389/fpsyg.2017.02062>
- [20] Lovallo, D., & Kahneman, D. (2003). Delusions of success. *Harvard business review*, 81(7), 56-63. <https://hbr.org/2003/07/delusions-of-success-how-optimism-undermines-executives-decisions>
- [21] MacDonald, G., Nail, P. R., & Harper, J. R. (2011). Do people use reverse psychology? An exploration of strategic self-anticonformity. *Social Influence*, 6 (1), 1-14. <https://doi.org/10.1080/15534510.2010.517282>
- [22] Mather, M.; Shafir, E.; Johnson, M.K. (2003). "Remembering chosen and assigned options" (PDF). *Memory & Cognition*. 31 (3): 422–433. <https://doi:10.3758/BF03194400>
- [23] Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 129-140. <https://doi.org/10.1177/2372732215600886>
- [24] Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: short-term benefits and long-term costs. *Journal of personality and social psychology*, 80(2), 340. <https://psycnet.apa.org/doi/10.1037/0022-3514.80.2.340>
- [25] Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of risk and uncertainty*, 1, 7-59. <https://doi.org/10.1007/BF00055564>
- [26] Sherbino, J., Kulasegaram, K., Howey, E., & Norman, G. (2014). Ineffectiveness of cognitive forcing strategies to reduce biases in diagnostic reasoning: a controlled trial. *Canadian Journal of Emergency Medicine*, 16(1), 34-40.pdf. <https://doi.org/10.2310/8000.2013.130860>
- [27] Soll, J. B., Milkman, K. L., & Payne, J. W. (2015). A user's guide to debiasing. *The Wiley Blackwell handbook of judgment and decision making*, 2, 924-951. <https://doi.org/10.1002/9781118468333.ch33>
- [28] Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- [29] Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157), 1124-1131. <https://doi.org/10.1126/science.185.4157.1124>
- [30] Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, 39(5), 806. <https://www.ucl.ac.uk/lagnado-lab/publications/harris/PsychRevInPress.pdf>

【研究ノート】**ChatGPT のデータ解析系授業への影響について**

三宅茂樹

Impact of ChatGPT on Data Analysis Courses

Shigeki Miyake

要旨：この研究ノートは、ChatGPT をデータ解析系授業に導入することの影響に焦点を当てている。主に、テーブルデータと画像データの演習問題に ChatGPT を使用し、その効果と限界を探求している。テーブルデータの問題では、ChatGPT がデータ整形からモデル評価までのプロセスを効率化し、平均的な精度のモデルを容易に構築できることが示されている。しかし、精度を向上させるには、データ解析の基本的な知識が不可欠である。画像データの問題では、ChatGPT の実行環境の制限によるボトルネックと、深層学習の基本的な知識の重要性が明らかになっている。全体として、ChatGPT は教育プロセスを効率化する可能性があるが、短期的には従来の教育方法の重要性は変わらない。一方、中長期的には授業の内容やアプローチの再評価が必要である。

1. 背景**1.1 ChatGPT について**

ChatGPT は OpenAI によって開発された先進的な自然言語処理モデルである。人間のよう流暢かつ自然な言語でのコミュニケーションが可能で、膨大なテキストデータから学習を行っている。このシステムは GPT (Generative Pre-trained Transformer) 技術に基づいており、大規模なデータセットからパターンを学習し、新しいテキスト生成に応用する。ChatGPT は質問への回答、テキスト生成、データ解析の課題解決など、幅広い用途に利用されている。特に、コンピュータプログラミングやデータサイエンス分野での応用が注目され、学生や専門家にとって重要なリソースとなっている。このモデルの主な特徴は柔軟性と広範な応用可能性にあり、日常のコミュニケーション、学問的な探求、技術的な問題解決において新たな可能性を提供している。

1.2 データ解析系授業への影響について

従来の授業設計では、最新の技術進化、特に ChatGPT のような高度な自然言語処理モデルの使用を想定していない。現代の技術環境において、このようなツールの導入は、学生が直面する実際の課題と授業内容の間にギャップを生じさせる。実際に ChatGPT を授業に用いた際の実践的報告が出始めているが、これらは主にプログラミングや語学教育に

関連している^{[1][2][3]}。データサイエンス系授業での具体的な実践例の報告はまだ少ない。[4]はデータ解析系授業に ChatGPT を導入した実践例といえるが、主に学生アンケート結果の集計について記述していて、導入の影響分析、対処法・改善案などの提案には至っていない。

ChatGPT には外部からデータを与えることでそのデータを解析する機能¹⁾があり、コーディングまで行うことが可能であるが、これは従来の授業設計の想定外である。そのため、授業内容や演習問題の調整に関しても、ChatGPT のようなツールを適切に組み込む方法については明確なガイドラインが存在しない。ChatGPT を授業に導入した場合のメリットやデメリットが未だ明確ではないが、教育者は授業の目的と手段を再評価し、どのような影響があり得るかを明らかにしながら、学生にとって最も有益な学習体験を提供するための新たなアプローチを模索する必要がある。

1.3 本研究ノートの役割

本研究ノートの主な役割は、ChatGPT を用いて典型的な演習問題に対して学生の立場でアプローチを行い、その過程で得られた知見を通じて、ChatGPT の可能性と限界を明確にすることにある。具体的には、ChatGPT がデータ解析の分野でどのように利用できるか、その機能と効果を詳細に分析する。また、学生が ChatGPT を使用する際のメリットとデメリットを検証し、教育的観点からその意義を探求する。さらに、授業への影響と、ChatGPT の導入に伴う教育方法の適応や改善策を提案する。なお本稿で使用した ChatGPT のバージョンは ChatGPT4 である。ChatGPT3.5 は無料で利用できるため、現時点で学生が最も使用しているバージョンであると予想できるが、大規模言語モデルの性能は時間が経てば各段に上がっていくため、本稿での調査や考察の急激な陳腐化を防ぐために敢えて 3.5 ではなく 4 を使用したことを注意しておく。

この研究ノートは、AI と教育の融合が進む現代において、教育者と学生の両方にとって重要な指針を提供することを目的としている。ChatGPT の持つ技術的な特性と教育への応用を深く理解することで、今後のデータ解析教育のあり方についての議論を深め、より効果的な学習方法の開発に寄与することを期待している。

2. ChatGPT を使った演習問題へのアプローチ

第2章では、ChatGPT のデータ解析系授業への影響を探るために、具体的な演習問題に焦点を当てる。特に ChatGPT を用いて典型的なデータ解析の演習問題を学生目線でのアプローチを行う。選択した問題は、データ解析コンペティションサイトである SIGNATE^[5]の練習問題から2題を取り上げた。SIGNATE はデータ解析のコンペを提供するサイトであり、その練習問題も実践的なデータ解析スキルの習得に役立つものが多い。本章では、これらの問題に対してどのように ChatGPT を使ってアプローチしたのかを記述する。特に、従来のアプローチと ChatGPT を用いたアプローチとの比較に重点を置き、その相違点や効果を分析する。この分析を通じて、ChatGPT がデータ解析の学習プロセスにどのように影響を与え、従来の教育方法にどのような新たな視点をもたらすのかを探求する。さらに、ChatGPT の使用が演習問題の解決方法にどのような変化をもたらすかについても詳

しく調査し、データ解析教育への具体的な応用可能性について考察する。

2.1 自動車の走行距離予測¹⁶⁾

2.1.1 問題概要

SIGNATE の練習問題「自動車の走行距離予測」は、典型的なテーブルデータを用いた問題であり、自動車のさまざまな属性からその燃費性能（1 ガロンあたりの走行距離）を予測することが目的である。この問題では、車種ごとの重量、排気量、馬力などの物理的特性に加え、年式や製造国などのデータが提供される。参加者はこれらの情報を基に、燃費性能を正確に予測するモデルを構築する。特に、本問題はデータ解析の基本的なスキルを試すのに適しており、回帰分析など機械学習アルゴリズムの理解と応用が中心となる。さらに重要な点は、データの処理や分析がすべて ChatGPT の実行環境上で行えることである。これにより、従来のデータ解析方法と比較して、ChatGPT を用いたアプローチがどのように異なるかを評価する絶好の機会を提供する。

2.1.2 ChatGPT を使ったアプローチ

まず、提供されたテーブルデータ（csv 形式）を ChatGPT に pandas dataframe として読み込ませ、以後データ解析の標準手順（データ整形、探索的データ分析（EDA）、モデル選択・学習、評価）に従った。

データ整形の段階では、欠損値の処理方法を ChatGPT に提案してもらい、その提案に基づいて欠損箇所を埋めた。次に、EDA では ChatGPT が目的変数と説明変数との相関をグラフ化し、データの関係性を明確に示した。この視覚化はデータの理解を深めるのに役立つが、従来はコードを手書きしながら行うのでコードのバグ取りもしながらデータの理解も進めなければならなかった。こういった状況はデータ解析を始めたばかりの者にとってはそれなりに難関であった。ChatGPT を使えばデータの理解に集中できるのが大きなメリットである。

EDA の結果を踏まえて、ChatGPT は回帰モデルを提案し、そのモデルで学習と評価を行った。また、ChatGPT はテストデータに対する予測実行から、予測結果を SIGNATE の投稿形式で保存するまで自然言語で指示するだけで完結するのも初心者には助かる。得られた結果は順位的に中程度またはやや上位であり、ベースラインとしては十分な結果である。これまでのプロセスには 10 分余りしかかからなかった。

ここから重要な点は、初回サイクルで得られたベースラインの結果からモデルをどのように改良していくか、である。この段階では利用者のデータ解析に関する知識（特徴エンジニアリング、学習モデルの選択、ハイパーパラメータのチューニングなど）が不可欠となる。ChatGPT は改良の方針を提案してくれるが、どの提案を選択するかは、利用者の経験と知識に依存する。

今回は学習モデルとして SVM を採用し、正則化係数とカーネルについてハイパーパラメータのチューニングを行った。ハイパーパラメータのチューニングにはこちらから指示しなくても scikit-learn の Grid SearchCV を使用していた。ハイパーパラメータのチューニングが完了した時点で投稿した結果は上位の 1 割程度であった。ここまでにかかった時間

はほぼ1時間であり、大変効率的にデータ解析のサイクルを回していることに気づく。

このように、ChatGPTはデータ解析のプロセスを大幅に効率化するが、最終的な判断や精度の改良には人間側の専門知識が重要な役割を果たすことが明らかとなった。

2.2 画像ラベリング (20種類)¹⁷⁾

2.2.1 問題概要

SIGNATEの練習問題「画像ラベリング (20種類)」は、前述の「自動車の走行距離予測」とは異なり、画像データを対象とする深層学習モデルの構築を目指す問題である。この課題では、動物、乗り物、自然の風景など20種類の異なるカテゴリーに属する画像が提供され、参加者はこれらを適切に分類するモデルを開発する。この問題は、テーブルデータを扱う前問題と異なり、コンピュータビジョンと画像処理の基礎知識が中心となり、特に畳み込みニューラルネットワーク (CNN) などの深層学習アルゴリズムの理解と応用が重要である。

この深層学習の問題は学習時に高い計算能力を要求し、特にGPUの使用が必須となる点が特筆すべきところである。ChatGPT内の実行環境ではGPUを使用できないため、Google Colaboratoryなどの別の実行環境を用意し、ChatGPTが提供するコードを実行する必要がある。提供された画像データセットを用いてモデルをトレーニングし、新たな画像に対する分類精度を高めることが求められる。この問題は、実際の画像分類の課題に近く、深層学習技術の応用能力を測る機会を提供する。テーブルデータと画像データ、それぞれの問題において必要とされるデータ解析のアプローチの違いを明確に示している。

2.2.2 ChatGPTを使ったアプローチ

「画像ラベリング (20種類)」問題へのChatGPTを使ったアプローチでは、まず訓練データをアップロードしたディレクトリを指定し、そこからデータを読み込むように指示した。しかし、最初に実行したコードでは画像データのサイズが大きすぎてメモリがクラッシュする問題に直面した。これに対処するため、画像データをバッチごとにメモリに読み込むようChatGPTに指示したところ、ジェネレータを用いたコードが回答された。このジェネレータを使用するために、訓練データが保存されているディレクトリ内にラベルに対応するサブディレクトリを作成する必要があったがChatGPTの回答に従ってこれを容易に実施できた。

学習モデルに関しては、最初からCNNを構築するのではなく、既存のモデルをファインチューニングする方針を取った。ちなみにChatGPTの提案はVGG16であった。モデルの学習後、テストデータに対する予測を行い、SIGNATEの投稿形式で保存するプロセスも前の問題と同様に実施した。

このようにベースラインの学習モデルまでは前問同様に容易に構築可能である。前問同様であるが本質的な問題はベースラインからどのようにモデルを改良していくかにある。今回は、data augmentation (データの増強) を行うように指示した。正確に述べると、ベースライン構築のコードはすでにdata augmentationを行う設定だったが初回のサイクルではオーバーフィッティングを起こすことが重要なので敢えてこれを行わなかった。

さらに、オーバーフィッティングを緩和する方法に関してもChatGPTより複数の回答

を得た。しかし、授業における演習を想定しているため、時間のかかるハイパーパラメータのチューニングまでは行わなかった。

結局、ChatGPT はデータの準備からモデルの評価に至るまでのプロセスを効率化する一方で、ベースラインからさらに精度を改善するためには深層学習の基本的な理解が依然として重要であることが理解できた。

3. 考察と結論

第2章で実施した SIGNATE 練習問題へのアプローチを通じて得られた重要な知見は、ChatGPT を活用することでデータ解析のプロセスが効率化される一方で、より高い精度を求める場合には従来同様専門的な知識が必須であることである。テーブルデータの解析では、ChatGPT の実行環境上でデータ整形から学習モデルの評価までの一連のプロセスを迅速に完了することが可能であった。これにより、従来の手法に比べてデータ解析のサイクルを格段に速く進めることができる。従ってグループワークの形式で問題に取り組む際には、作業単位での役割分担の意義が薄れ、代わりにデータ理解やアイデアの交換がより重要な役割を果たすようになると考えられる。

一方で、GPU を要する深層学習の問題においては、ChatGPT の実行環境だけでは不十分であり、Google Colaboratory などの外部環境との連携が必要となる。従って、学習時間が作業のボトルネックとなる点は従来と変わらない。テーブルデータの問題と同様に、コードやアルゴリズムに詳しくない場合でも平均レベルのアウトプットを容易に得ることができるが、アウトプットの精度をさらに高めるためには深層学習の深い理解が必要となる。

これらの知見は、ChatGPT の利用がデータ解析の教育や実務に与える影響を考える上で重要である。一方で効率化が進み、もう一方で高度な成果を得るためには、従来と同様に深い専門知識が求められる。このバランスを理解し、適切に対応することが、今後のデータ解析教育や実務の発展において鍵となる。

3.1 メリットとデメリット

学生が ChatGPT を活用して演習問題に取り組むことには、明確なメリットとデメリットが存在する。メリットとしては、まずコーディングに関する手間が軽減され、分析そのものに集中できる点が挙げられる。特にテーブルデータの解析など、ChatGPT の実行環境内で完結できるケースでは、データ解析のサイクルを大幅に速めることが可能となる。これにより、データの深い理解や学習モデルの改良に向けたアイデアの深掘りにより多くの時間を割くことができる。

しかし、一方でデメリットも存在する。ChatGPT を用いることで、自然言語で指示を出すだけで適切なコードが返ってくるため、学生がコーディングスキルを習得する機会が減少する可能性がある。これは、コードを記述する技術やプログラミングの論理的思考を養う過程が疎かになるということを意味する。プログラミングスキルはデータ解析だけでなく、様々な分野での問題解決や新たなアイデアの実現に不可欠であるため、このスキルの習得は重要である。

3.2 授業への影響

ChatGPT の利用がデータ解析の教育に与える影響を考えると、コードやアルゴリズムの理解が依然として重要であることが明確になった。本節では今後1～2年スパンでの短期的な授業への影響と、それよりも長いスパンでの中長期的な授業への影響について考える。

3.2.1 短期的な影響

プログラミングの基礎知識やデータ解析の論理的な思考プロセスは、ChatGPT を用いても変わらず重要な要素である。従って、授業の大枠に関しては、従来のアプローチを維持する必要がある。ただし、演習問題に ChatGPT を取り入れることで解析のサイクルを数倍速めることが可能となる。これにより、学生はより迅速にそれなりの結果を得られるようになる。この変化は、授業における焦点を学習モデルの初期構築から、より高度な改良や最適化のプロセスへとシフトさせる。教師は、この新しい動向に対応するために、学生がモデルを深く理解し、効果的に改良する方法について指導する準備を整える必要がある。

総じて、ChatGPT の導入は短期的には教育の方法を根本から変えるものではないが、教師と学生の両方に新たな機会と課題を提供する。このツールを効果的に活用し、同時にデータ解析の基本的なスキルと理解を維持・深化させることが、今後のデータ解析教育の発展に不可欠である。

3.2.2 中長期的な影響

中長期的な観点から ChatGPT の導入が教育に与える影響を考えると、その効果はデータ解析の授業に限定されない。ChatGPT を活用することで、対話形式の自学自習が可能となり、学習者はデータ解析をはじめとする多様な分野で知識を深めることができるようになる²⁾。これにより、自律的な学習スタイルが促進される。しかし、ChatGPT は現時点ではハルシネーションの問題を抱えており、完全に信頼できるチューターや情報源として機能するには至っていない。ただし、技術の進歩によりこの問題は遠くない将来には大幅に緩和される可能性がある。

近い将来、学生は ChatGPT を使って高度な概念も自学自習できるようになり、かつ、ChatGPT の解析手法もベースライン以上に進むことができるレベルに到達すると予測される。その結果、大学の授業形態自体にも変化が生じる可能性がある。基本的な概念の習得は学生自身の自学自習に委ねられ、大学の授業では概念の習得の確認や応用的な問題の演習に焦点が置かれるようになると考えられる。これは、反転授業や分野をまたいだプロジェクトベースの演習が主流となることを意味する。

中長期的な影響の文脈において、データ解析系授業への具体的な影響は以下のように考えられる。まず、反転授業においては、学生が ChatGPT を活用して教師が与えた問題を解析し、その結果をグループワークで討議する形式が一般的となると予想される。ChatGPT のアプローチは単一ではないと予想される³⁾ので複数の結果を持ち寄ることで、学生はベースラインからどのような工夫を加えることで精度を向上させることができたのかを深く理解し、共有する機会を得る。

プロジェクトベースの演習については、学生が自らテーマを探求し、それについての解

析を行うという基本的な枠組みは現在と変わらない。ただし、ChatGPT との対話を通じて、全ての参加学生がベースライン以上の精度に至ることが可能となる。これにより、各学生のプロジェクトへの関わりが従来よりも深まり、より充実した学習経験を提供することになる。

中長期的には、ChatGPT の利用が教育全体のあり方に影響を及ぼす可能性がある。データ解析系授業においても、基本的な概念の習得は学生自身の自学自習によって行われ、大学の授業ではその確認や応用的な問題の演習に重点を置く形式が主流となる見込みである。このように、ChatGPT は学習方法や授業形式に変革をもたらし、学生が体系的な知識と批判的思考法を効果的に身に着けるための新たな手段を提供することになると考えられる。

注

- 1) 当初 "Code Interpreter" という名称で導入されたが、その後 "Advanced Data Analysis" に改名され、現時点では特定の名前はあてがわれていないようである。
- 2) プログラミング言語の自学自習の例は既に教科書レベルで紹介されている^[8]。
- 3) 現時点でも例えばベースラインからの向上案として特徴エンジニアリングやハイパーパラメータのチューニングなど複数の案が提示される。

参考文献

- [1] 倉光君郎, 「ChatGPT と Colab の連携: ChatGPT を活用したプログラミング演習の実例」, 教育機関 DX シンポジウム (2023 年 5 月 12 日), available at <https://edx.nii.ac.jp/lecture/20230512-06>.
- [2] 柳瀬陽介, 「大学英語教育における ChatGPT 活用型授業実践: 英語教師が認識する生成系 AI 活用の可能性と限界」, 教育機関 DX シンポジウム (2023 年 10 月 13 日), available at <https://edx.nii.ac.jp/lecture/20231013-05>.
- [3] 山田優, 「大規模言語モデル ChatGPT を活用した翻訳トレーニングと外国語教育」, 教育機関 DX シンポジウム (2023 年 11 月 13 日), available at <https://edx.nii.ac.jp/lecture/20231113-06>.
- [4] Y. Zheng, "ChatGPT for teaching and learning: an experience from data science education," available at arXiv:2307.1665v1 [cs.CY], 2023.
- [5] <https://signate.jp/>
- [6] 自動車の走行距離予測, available at <https://signate.jp/competitions/121>.
- [7] 画像ラベリング (20 種類), available at <https://signate.jp/competitions/108>.
- [8] 小野哲, 「ソフトウェア開発に ChatGPT は使えるのか? - 設計からコーディングまで AI の限界を探る」, 技術評論社, 2023.

【研究ノート】

STEAM 教育に適した 参加度向上のための環境開発

兒玉賢史・爰川知宏・武本充治・

辻野雅之・藤井竜也・山本裕

Development of an environment suitable for STEAM education to improve participation

Satoshi Kodama, Tomohiro Kokogawa, Michiharu Takemoto, Masayuki Tsujino,
Tatsuya Fujii and Hiroshi Yamamoto

Abstract : Due to the worldwide outbreak of the new coronavirus disease 2019, many lectures had to transition from the face-to-face to distance learning format. Consequently, new and different ways of delivering lectures have been explored, including via the PC or tablet. Currently, many lectures are returning to the traditional face-to-face format. However, more effective learning methods are being sought by incorporating some of the lecture methods that have been used in distance classes into face-to-face classes. Therefore, we decided to verify whether we can create an environment for more active student participation in lectures by incorporating various lecture methods utilized in distance classes into face-to-face classes at our university. As it is in the process of being verified, this report describes the current status of the construction of the verification environment to select the functions that are needed.

Keywords : teaching effectiveness, higher education, active learning, utilization of ICT, Learning and learning models, STEAM

1. まえがき

新型コロナウイルスの世界的大流行により、多くの講義が対面授業から遠隔授業となっていた [1]。そのため、PC やタブレットを利用して受講するというこれまでとは異なった新たな講義が模索されてきた [2]。現在では、感染症法上の位置付けが第 5 類感染症に移行したこともあり、外出の自粛要請や行動制限等が撤廃されたことから、講義の多くが従来の対面授業へと戻りつつある [3, 4]。その一方で、アフターコロナ時代を見据えた新たな STEAM 教育による教科横断型教育への対応や、アクティブラーニングにおける受講者の能動的な学習法も求められている [5-8]。また、より学習効果の高い講義を行うために、遠隔授業で行われてきた講義スタイルの利点を対面授業にも取り入れるといった研究もな

されている [9, 10]。そこで本学においても遠隔授業における様々な講義法を対面授業に取り入れることで、より積極的に講義に参加するための環境が構築できるか検証を行うこととした。

最終的な評価を得るためには、一定期間、複数の科目において有効性を確認する必要があるため、本稿では、途中経過として、必要と思われる機能の検討と検証環境の構築状況について記述することとする。

2. 研究目的

遠隔授業は、単純に PC 等を通して講義を一方的に受講するだけでなく、必要に応じてノートや小テスト等のデータをすぐにアップロードすることが可能である。また、講義内において質問等を容易に行うことができるといった特徴を持つため、対面よりも情報共有がしやすいといった評価がある [8, 11]。さらに細かい画像やグラフ、多くの情報を講義内に提示できるといったメリットもあることから、対面授業においても遠隔授業を応用することによって、学習効果を高めるための研究が行われている [8, 9, 11, 12]。そこで、本学においても STEAM 教育やアクティブラーニングの一環として、従来の対面授業よりも積極的に講義に参加できるか検証を行うこととした。

最終的な評価を得るには、対象となる科目の選定と一定期間の調査が必要であるため、2022 年度（本年度）は、有効と思われる機能の最小環境を構築することとした。

3. 環境構築の内容

講義への参加度や学習意欲を向上させるために、対面授業においても有効であると考えられる機能を選択した。今回、特に検証する機能として、「学習記録のアップロード」、「課題等を個別配信するための Slack の利用」、「複数台のカメラによるストリーミング配信」、「チャット機能」、および、「アクティブウインドウのチェック」といった環境について確認を行うこととした。本年度は初期実験として、機能そのものの有効性を検証する必要があることから、評価するための最低限度のシステム開発を行った。

3.1 学習記録のアップロード

一般的な座学では、どうしても一方向性で受け身型の講義になってしまうため、適宜、問題等を解決して結果を提出してもらう必要がある。しかしながら、広く使われている Learning Management System (LMS) は、レポート等をアップロードできるものの、個々の学習管理を目的としているため、講義中に全体の理解度や参加度合いをその場でチェックすることには適していない。そこで、近年では、学生の多くがスマートフォンを携帯していることから、記述したノートやプログラミングの出力結果等をカメラ撮影により教員側で即座に確認できるシステムを構築した (図 1, 図 2)。本機能は、教員側のサーバに画像ファイルとして即時反映されるため、現在の解答状況や途中式の確認等が簡単にできると考えられる。なお、Web ブラウザを介して撮影とアップロードが行えるようにすることで、学生側も容易に提出できるように開発した。

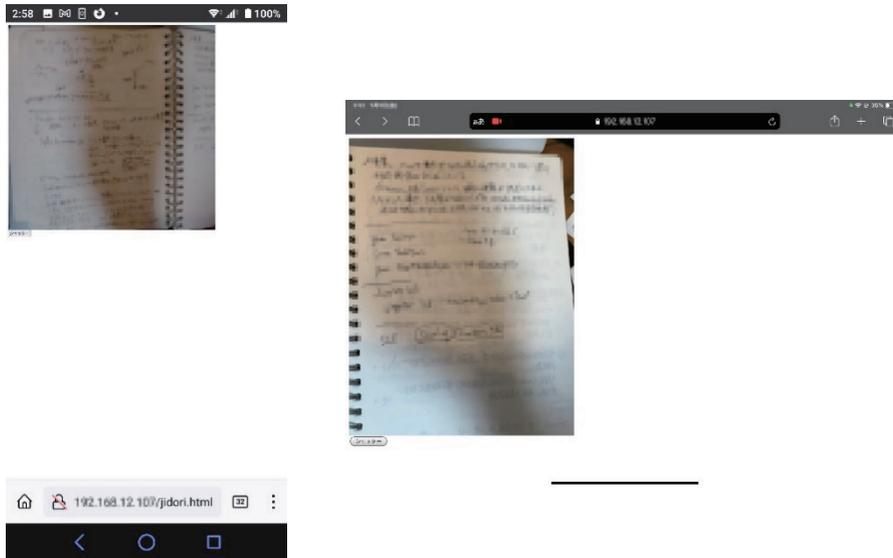


図 1. スマートフォン、および、タブレットからの撮影とアップロード

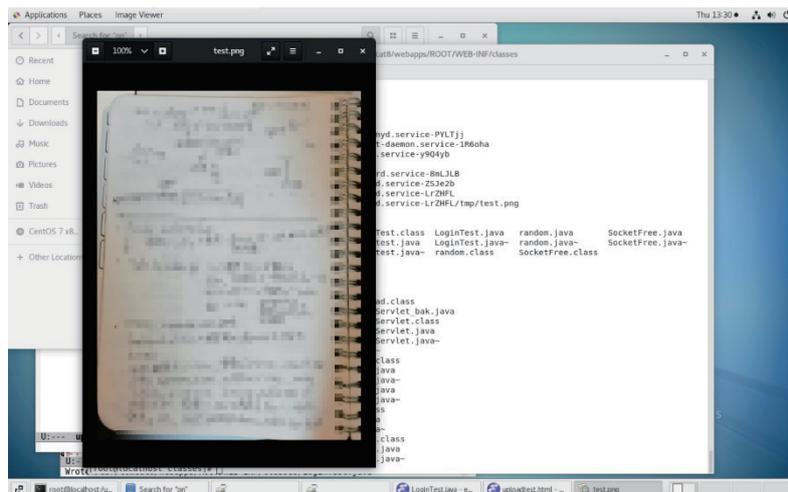


図 2. 撮影結果の閲覧

3.2 個別の問題や資料等の Slack 配信

本学学生は、各自 Slack のアカウントを持っており、講義の連絡等にも利用している。Slack は、学生個人やグループに対して、文字列だけでなく、データの受け渡しも可能であるため、学生に応じた問題や資料等の個別配信が行いやすいと考えられる。そこで、講義の進行速度に応じて、バッチ処理により、効率よく問題や資料の個別配布が行える機能を用意した。Slack は API を提供しており [13]、それらを利用することで効率よく送信ができる環境を構築した (図 3.)。

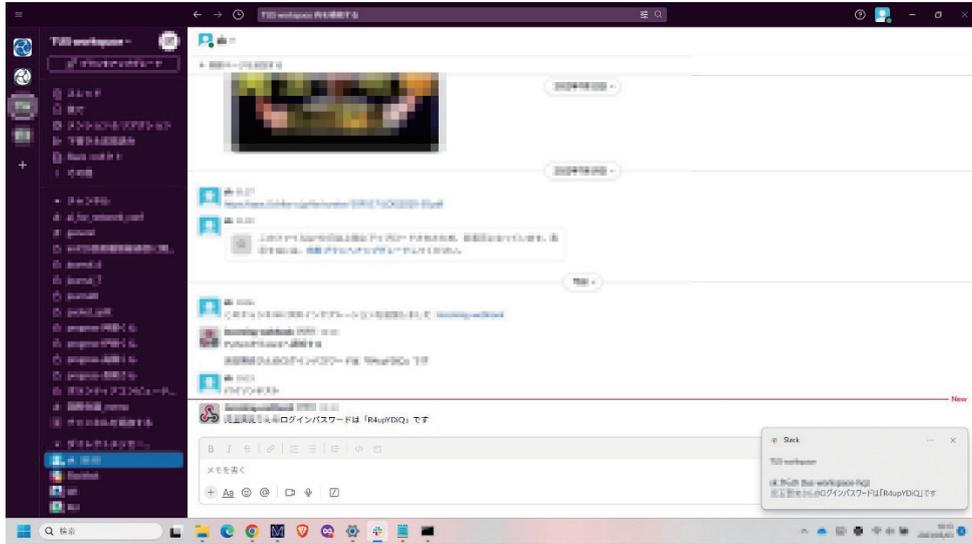


図 3. バッチ処理による Slack への個別投稿

3.3 複数台のカメラによるストリーミング配信

プリント基板上の端子やシングルボードコンピュータ上の汎用入出力端子等は非常に小さいため、講義中に接続状態や LED の発光状態等を正確に伝えることが難しい。そのため、遠隔授業で用いた書画カメラや配信法等を用いて表示できる機能は有用だと考えられる (図 4)。しかしながら、学習速度は受講者毎に異なるため、複数の場面を配信する場合、適切なタイミングで切り替えることが難しい。そこで、全体的な映像と詳細な端子の位置といった複数の映像 (複数台のカメラ映像) を同時に配信できるシステムを作成した。本年度は、配信方法の有効性を検証ための環境構築が主目的であることから、汎用性のあるブラウザや VLC media player 等で閲覧できる MJPG-Streamer [14, 15] を用いて環境を構築した (図 5)。本手法では、複数のポートを用いて、同時にストリーミング配信ができるため、各受講生のディスプレイ上の複数のウインドウを各々の適したタイミングで切り替えられることから、各自の学習スピードで確認できると考えられる (図 6)。



図 4. 基板の拡大表示

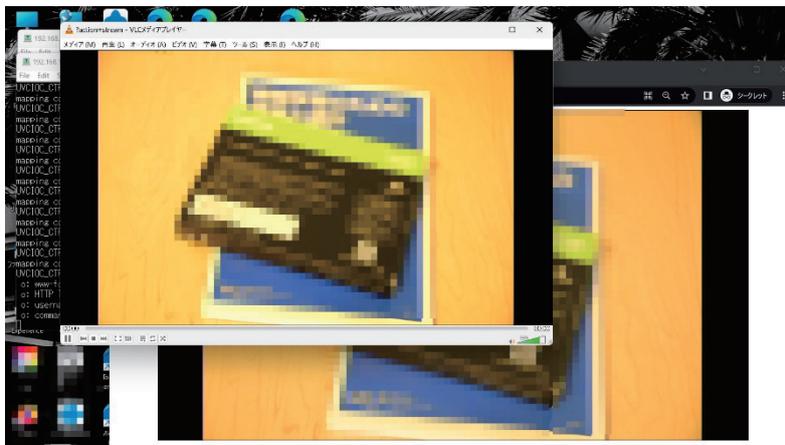


図 5. ブラウザや VLC media player による閲覧

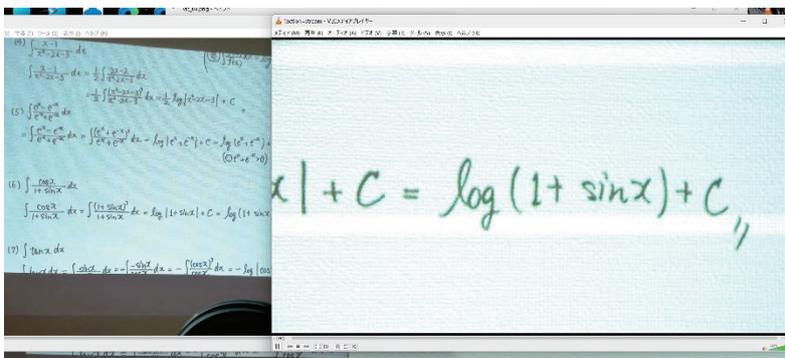


図 6. 全体像と拡大表示の同時配信

3.4 チャット機能

前述したように、遠隔授業の方が質問を行いやすいとの評価がある [8, 11]。そこで、簡易的に講義内において質問を行うための機能を用意した。対面授業で運用するため、本機能は、主に学生側から質問を受け、教員側が口頭で説明するために利用することを想定しているが、直接全体に返答を行うことも可能となっている (図 7.)。

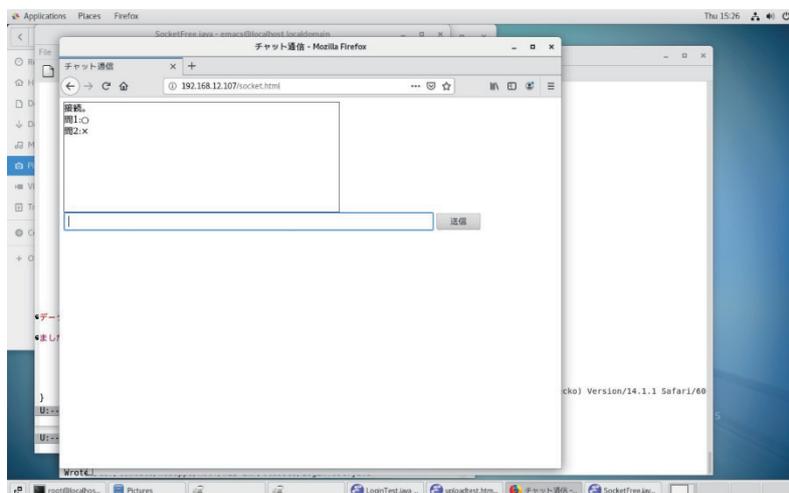


図 7. 質問受付用のチャット機能

3.5 アクティブウィンドウのチェック

現在、どのウィンドウがアクティブになっているのかを確認して、ログとして取得できるシステムを構築した。本機能を利用することで、どのウィンドウに受講生が着目しているのかといった情報や、教員側が講義している内容に連動して、受講生のウィンドウがどのように変化しているのかといった記録が取得できる(図 8)。本システムは、作業内容の推定を目的としたログパターン解析 [16, 17] と同様であることから、積極的に講義に参加しているかといった状態も確認可能であると考えられる。ただし、本機能は各ウィンドウのタイトルを取得できるという性質上、個人情報に繋がる情報も収集できてしまうことから、プライバシー等に配慮する必要があると思われる。

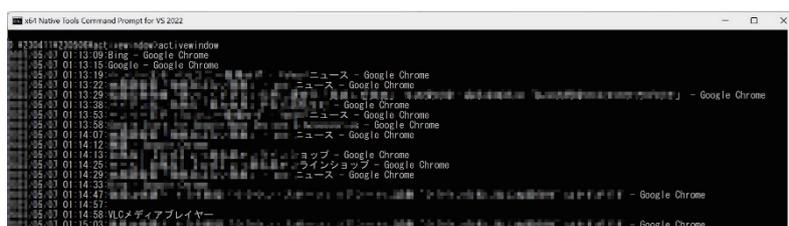


図 8. アクティブウィンドウによる受講生の閲覧状況の確認

4. まとめ

遠隔授業は対面授業と異なり、学生全体の反応をその場の雰囲気を得ることは難しいが、個々の学生からの直接的な反応や現在の理解度を確認しやすいといった長所がある。そのため、遠隔授業で用いた講義法は、対面授業においても利用できる部分は多いと思われる。本年度は主に対面授業においても受講生が積極的に学習するために必要となる機能を

選定するとともに、それに基づいて有用と思われる検証環境を構築した。次年度は、実際に学生からの反応を得ることで、各機能の有効性を検証していくとともに、引き続きSTEAM教育やアクティブラーニングに応用可能と考えられる機能の改善と環境構築を行いたい。

参考文献

- [1] 村上 正行, コロナ禍がもたらす大学教育の可能性 — 対象・方法・内容 —, 大学教育学会誌, 44 (1), pp. 76-77, 2022.
- [2] 磯 直行, システムはハード・ソフト・ネットの組合せ 先を見越した設計が功を奏す, 中京大学, https://www.chukyo-u.ac.jp/research_2/news/2023/01/021683.html. (アクセス日:2023年3月15日)
- [3] 大学等における新型コロナウイルス感染症への対応状況について, 文部科学省, https://www.mext.go.jp/content/20200917-mxt_koutou01-000009971_14.pdf, 2022. (アクセス日:2023年1月10日)
- [4] 折戸 洋子, 崔 英靖, 岡本 隆, 岡本 直之, 曾我 亘由, 橘 恵昭, 新型コロナウイルス感染症による大学生活への影響: 大学生は Before コロナに戻ることができるのか?, 愛媛大学社会共創学部紀要, 第7巻 (1), pp. 11-29, 2023.
- [5] 菊澤 育代, イノベーション力を育む多様な学び — ICT教育、STEAM教育、デザイン思考教育の考察を通して —, 都市政策研究, 第22号, 2021.
- [6] 大橋 裕太郎, 山地 秀美, 糸野 文洋, 辻村 泰寛, コロナ禍の高等教育でのアクティブラーニング科目の運営と変化, 工学教育, 70 (3), 2022.
- [7] 大島 まり, STEAM教育への取り組み, 中央教育審議会初等中等教育分科会教育課程部会, 第120回, https://www.mext.go.jp/content/20200917-mxt_kyoiku01-000009959_4.pdf, 2020. (アクセス日:2023年1月7日)
- [8] 岡田 佳子, 学生からみたオンライン授業のメリットとデメリット —オンライン環境下のアクティブラーニングに焦点を当てて—, 長崎大学教育開発推進機構紀要, 第11号, 2021.
- [9] 岡島 寛, アフターコロナの大学授業 (スライドか? 板書か? 国立大学工学部), https://researchmap.jp/blogs/blog_entries/view/390442/864fa15a7765f76f43ebe0595a2a3953?frame_id=848894, 2022. (アクセス日:2023年1月7日)
- [10] 鈴木 克明, 大学教育の新たなブレンド型モデルの構築に向けた提言, 教育システム 情報学会, 第46回全国大会, pp. 77-78, 2021.
- [11] 遠隔授業に関するアンケート調査の概要, 国立情報学研究所, https://www.nii.ac.jp/event/upload/20200914_Report.pdf, 2020. (アクセス日:2022年12月20日)
- [12] 橋本 和幸, コロナ禍にオンライン授業と対面授業の両方を受講した学生による授業評価, 徳寺大学研究紀要 (16), pp. 137-150, 2022.
- [13] Introduction to Slack apps, <https://api.slack.com/start/apps>. (アクセス日:2022年12月20日)
- [14] G. Sumalatha, S. Bharathiraja, Implementation of Real Time Video Streamer System in Cloud, International Journal of Engineering and Applied Sciences, vol. 3(4), 2016.
- [15] S. N. Kale, R. D. Patil, Remote Video Monitoring System Using Raspberry Pi 3 and GPRS Module, International Journal on Recent and Innovation Trends in Computing and Communication, vol. 5(10), 2017.
- [16] 宇佐美雄基, 石沢千佳子, 景山陽一, PC上で行われた作業内容推定を目的としたログパターンの解析, 情報処理学会全国大会講演論文集, 80 (4), pp. 663-664, 2018.
- [17] MylogStar, RUNEXY, https://www.mylogstar.net/images/casestudy/jindai/mylogstar_jindai.pdf. (アクセス日:2023年5月12日)

【研究ノート】

「ヒト・モノ協調型のスマートモビリティ社会」 に関する研究報告

辻野雅之・内埜大斗・水上憲明・山口直彦

Research Report on “ Smart Mobility Society based on Human-Object Coordination ”

Masayuki Tsujino, Hiroto Uchino, Noriaki Mizukami, and
Naohiko Yamaguchi

Abstract: We aim to realize a “ smart mobility society based on human-object coordination” within the framework of the envisioned smart society in Society 5.0. This paper presents the current status of our research efforts, with a focus on simulating transportation planning for online demand-responsive transport (DRT) as an essential aspect of the research aimed at achieving our goals.

Keywords: smart mobility society, online demand-responsive transport (DRT), numerical simulation

1 はじめに

Society5.0 [1] が目指すスマート社会の到来を見据えつつ、「ヒト・モノ協調型のスマートモビリティ社会」という構想のもと、その概念の有効性検証を行っている [2]。そこでは、自動運転車・自律飛行 Drone のような自律的に移動する機器・装置（“AMO：Autonomous Mobile Object”）が重要な役割を果たす。社会システム全体での視点で、AMO を効率的に制御する方法が課題となる。このため、我々は、広域に展開された 5G による通信環境、高度化された IoT 技術、クラウド・エッジのサーバ技術を利用して、図 1 に示す 2 つの異なるタイムスパン（秒未満・分以上）で AMO を制御することで管理対象の社会システムを最適化することに関心がある¹⁾。

本稿では、この「ヒト・モノ協調型のスマートモビリティ社会」の研究状況の説明として、都市レベルでの最適化を狙いとする「オンデマンド交通の輸送計画シミュレーション」を中心に検討状況を示す。

1) 図 1 では、AMO に加え、本構想で制御対象とするヒトが操作する装置（図中、“HDO：Human-Driven Object”）も記載しているが、本稿の本文では HDO には言及しない。

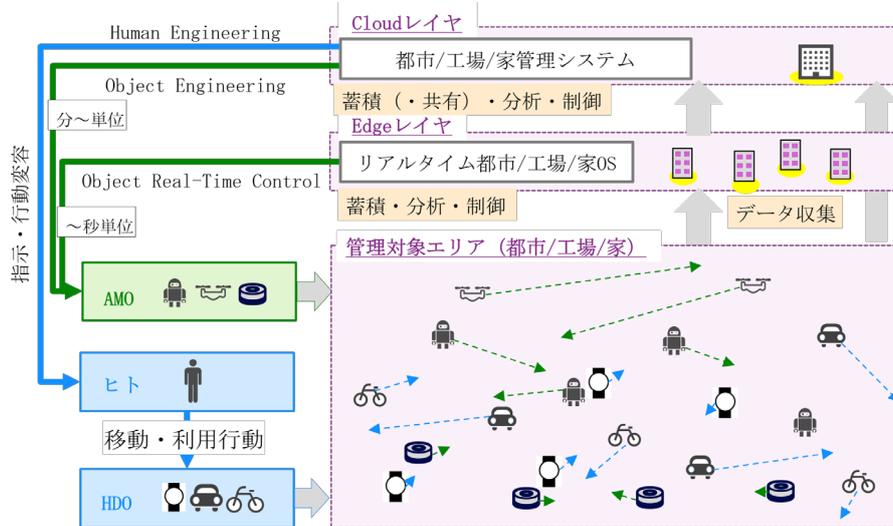


図1. システムイメージ

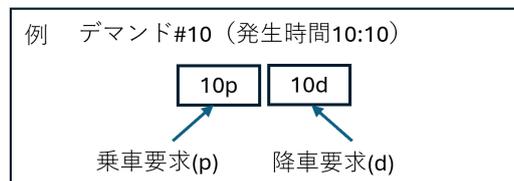
2 輸送計画シミュレーション検討状況

本章では、交通流シミュレータのオープンソースソフトである Simulation of Urban Mobility (SUMO) [3] を用いて行った輸送計画法の数値実験評価について説明する。

2.1 オンデマンド交通の輸送計画

マイカーに代わる高齢者向けの近距離移動手段として、オンデマンド交通が注目を集めている [4]。オンデマンド交通は、利用者からの要求（デマンド）に応じて乗り合いでの輸送を行うサービスである。随時発生するデマンドでは、利用者が乗車地点・降車地点を指定する（図2 参考）。輸送車には、デマンドの乗車地点まで利用者を迎えに行き、そこで利用者を乗せて、降車地点まで送り届けることが求められる。

デマンドは随時発生する
デマンドでは乗車地点・降車地点が指定される



※乗車要求・降車要求は地点を含む
(説明では、乗降順序のみに言及するので記載省略)

図2. デマンドの構成

このサービスを、複数の輸送車と、それらに対して以下の輸送計画・経路計画に沿った割当を行う制御装置からなるシステムで提供する。

輸送計画 各利用者を輸送する輸送車の割当、及び、各輸送車で利用者の乗降順序割当を行う。

経路計画 輸送車が、利用者の乗車・降車する地点間を移動する経路を選択する。

本研究では、経路計画は道路網の構成から自動的に求められる地点間の最短経路を選択するものとし、システムで発生するデマンドをどの輸送車に割当て、乗車・降車地点をどのような順序（乗降順序）で走査するかという輸送計画を対象とする。

輸送計画に対しては、以下の前提を置く（図3参考）。

- 新規にデマンドが発生した時点で輸送計画を実行する（策定された計画に基づき割当を行う）。
- 既に輸送車に乗っている利用者は、その利用者が指定する降車地点まで送り届ける（利用者を途中で降ろし、他の輸送車へ乗り換えさせない）。
- デマンド発生から所定の時間迄に乗車地点に利用者を迎えに行く必要がある（所定時間内に迎えにいけないデマンドはキャンセルされる）。
- 各輸送車はその定員を超えた人数の利用者を乗せられない。
- 未発生デマンドに対する情報を持たない（そのため、将来発注するデマンドの内容を踏まえた輸送計画は策定できない）。

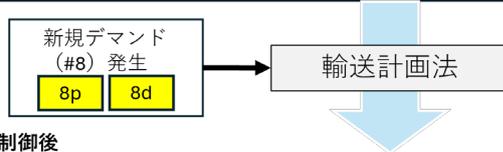
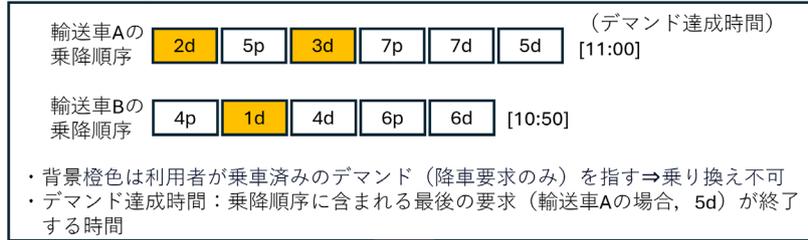
この前提の上で、以下に示す4つの輸送計画法（従来法 [ex1]、提案法 [ex2-4]）を比較評価対象とする。

ex1：逐次最適挿入法（従来法） 逐次最適挿入法 [5] では、既存のデマンドに対しては、輸送車／乗降順序割当の見直し（再割当）を行わない。その上で、新規デマンドを受け入れるために、その乗車・降車要求を挿入する既存の乗降順序での最適な位置を探索する。これは、全ての挿入可能な位置について挿入後のデマンド達成時間を計算し、そのデマンド達成時間が最も早い位置に挿入して新たな輸送計画を策定する（図4参考）。

この逐次最適挿入法では、各輸送車への利用者の再割当を行わず、既存の乗降順序の挿入位置に新規デマンドの乗降位置を割当てただけである。このため、輸送車内に閉じた割当と捉えることができる。さらに良い割当とするためには、この割当の範囲を拡張することが有効である。そこで、未乗車の利用者に対して割当てた輸送車を変更するよう輸送車間へ割当の範囲を拡張する。

図5で示すように、新規デマンドの発生時に拡張範囲とする輸送車を選択し、それらの輸送車に割当済みのデマンドを1つ選び、その選定デマンドを新規デマンド同様に逐次最適挿入法で割当てし直す「輸送車間再割当法」を提案する。逐次最適挿入法では、どのように新規デマンドが発生しても、一旦、輸送車に割当てた利用者を変更することはなかった。提案法では、新規デマンド発生時点で随時この再割当を行うことで、輸送計画が隘路に入り込むのを防ぐ効果を狙いとしている。

制御（割当・選択）前



制御後

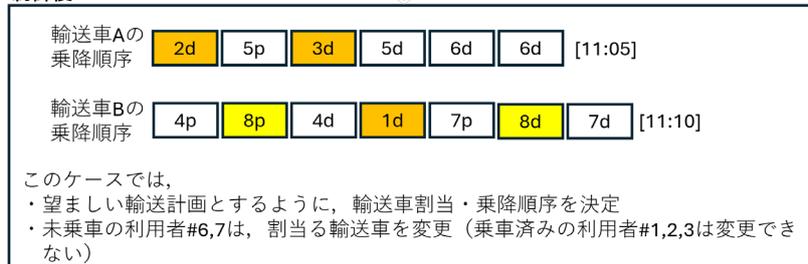
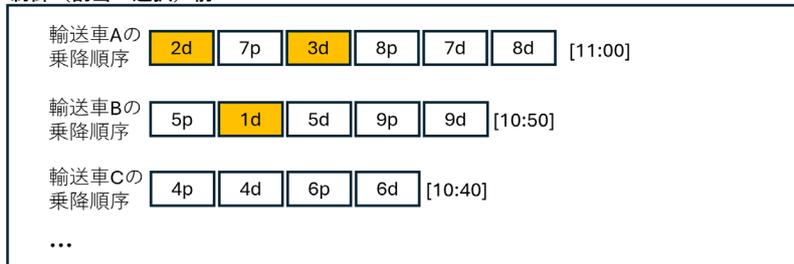


図 3. 輸送計画

制御（割当・選択）前



逐次最適挿入法:ex1

Step1. 新規デマンドの乗車／降車要求を既存の乗降順序の挿入可能な全ての位置に挿入する。なお、ここでの挿入可能性を定める要因は、輸送車の定員を超えないか、所定時間内に新規デマンドで指定する乗車地点に迎えに行けるか、及び、降車要求に先駆けて乗車要求に応えられていることである。その際、新規デマンド挿入後の乗降順序のデマンド達成時間を計算する。



Step2. 新規デマンド挿入後に計算した デマンド達成時間が最も早い挿入位置 に基づき、乗降順序を変更する。

制御後

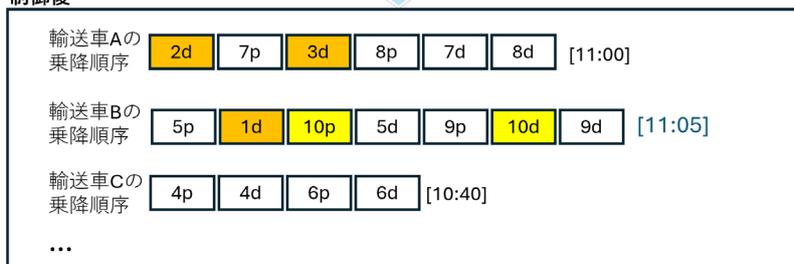
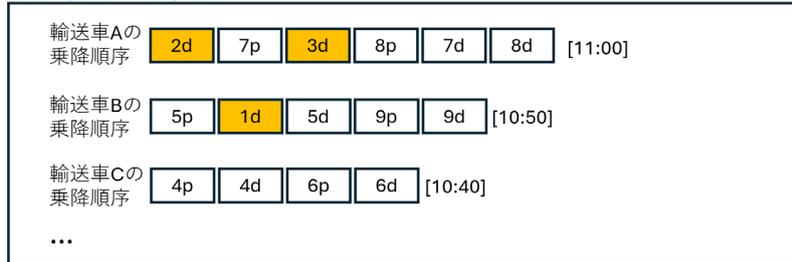


図 4. 逐次最適挿入法

制御（割当・選択）前



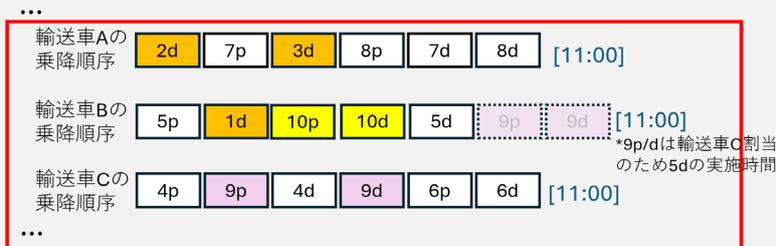
輸送車間再割当法:ex2-4

Step1. 逐次挿入最適法で新規デマンドを割当る輸送車を含む複数台の輸送車を選定する（選定法が、ex2-4で異なる部分）。
 ⇒以下に示す例では、輸送車A～Cが選定されている。
Step2. 選定輸送車に割当られているデマンドより、再割当可能なデマンドの内、まだ再割当の対象となっていないデマンドを選ぶ。デマンドが選択できたときはStep3に、再割当対象となっていないデマンドが無いときはStep4に移る。



7p 7d 再割当対象デマンド (#7)

Step3. 再割当対象デマンドと新規デマンドの乗車／降車要求を順に、選定輸送車の全ての挿入可能な位置に挿入する。その際、新規デマンド挿入後の乗降順序のデマンド達成時間を計算する。終了したら、Step2に移る。



Step4. 選定輸送車の最遅のデマンド達成時間が最も早い挿入位置に基づき、乗降順序を変更する。

制御後

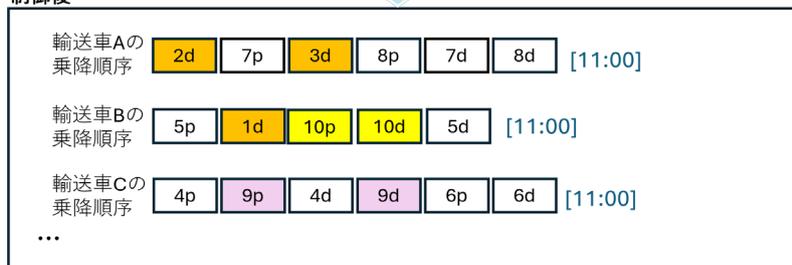


図 5. 輸送車間再割当法

割当範囲の拡張はより良い割当が求まる可能性が得られる反面、走査する輸送車の台数が増えたときに輸送計画に要する計算負荷が大きくなるという問題がある。そこで、計算時間を短縮させるため、再割当対象とする輸送車を絞り込む（ex3,4）ことを検討する。輸送車間再割当法として、再割当対象とする輸送車の選定法に応じた以下の3つの方法を比較評価対象とする。

- ex2：輸送車間再割当法（全輸送車選定）システム上の全輸送車を対象とする。輸送車の台数によっては、計算負荷が課題となる。
- ex3：輸送車間再割当法（逐次最適挿入法選定）逐次最適挿入法を行い、後述するデマンド達成時間の評価尺度で上位に評価される輸送車を目標とする台数分選定する。
- ex4：輸送車間再割当法（ランダム選定）全ての輸送車から、目標とする台数分の輸送車をランダムに選定する。

2.2 評価条件

本節では、SUMO を用いて行ったシミュレーションの条件について説明する。図6に示す、各辺 200m で格子状に構成された道路網上を5台の輸送車が走査する。各輸送車の定員を4人とする。乗車人数を1人とするデマンドが、60秒毎にシステムで発生する。なお、デマンド発生から1,800秒（30分）以内に乗車地点に利用者を迎えにいけない場合、そのデマンドはキャンセルされる。本来、デマンド発生次第、輸送計画法に基づく割当制御を行うことを想定しているが、本シミュレーションでは実装の便宜上30秒間隔で周期的に新たに発生したデマンドに対する輸送計画を実施する。なお、この輸送計画実施の周期は、シミュレーション条件としている60秒間隔のデマンド発生頻度より短いため、一度の周期で割当制御対象となるデマンドは最大でも1件である。

輸送計画の評価であることを意識し、道路網には本システムで対象とする輸送車（5台）しか存在しない、また、交差点には信号を設置しない。ex3,4では、それぞれの方法に則り、再割当対象の輸送車を5台から3台に絞り込む。

その上で、実際の道路網での状況を模した評価とすべく、(1) 輸送車の速度は一定では

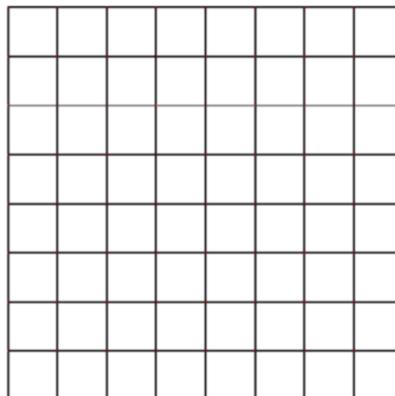


図6. シミュレーション対象の道路網

なく、所定の最高速度（30km/時）を守りつつ交差点で加減速を行う、(2) 輸送車は、利用者の乗降時に 60 秒間停車する、という走行ルールを適用する。

2.3 評価結果

本節では、シミュレーションで行った、4つの輸送計画法の評価結果について説明する。前節で示す条件のもと、シミュレーション時間を 10,000 秒とし、この間で 60 秒間隔での 150 件のデマンドを発生（最後の 1,000 秒はデマンド未発生）させ、評価対象とする ex1-4 の輸送計画法を適用した場合のシミュレーションを行った。なお、シミュレーションでは、以下の評価指標に基づき、輸送計画法を評価した。

処理時間 (Handling Time) シミュレーション実行から終了までにかかる時間。新規デマンドに対して輸送計画策定による割当制御に要する処理時間の概算に活用。

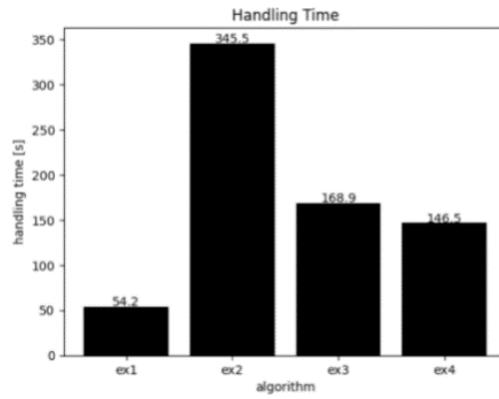
デマンド達成率 (Demand Achievement Rate) 発生したデマンドの内、シミュレーション時間内に、降車地点に送り届けられたデマンドの割合。キャンセルされたデマンド、時間切れになったデマンドがどのくらい少ないかを示す指標。

平均デマンド完了時間 (Average completion Time) デマンドが発生してから降車地点へ送り届けるまでに要する時間の平均。

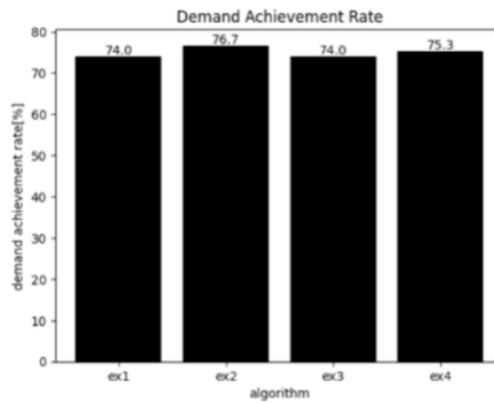
これまでに説明した条件によるシミュレーションの評価結果を図 7 に示す。この図は、10,000 時間のシミュレーションを 10 回実施した時の評価指標の平均値を示す。

図 7(a) で示す処理時間は、150 件のデマンドに対する輸送計画法の実行、10,000 時間のシミュレーションの実行、評価指標算出に関する計算時間を含む。この計算時間の大部分が輸送計画実行に要する計算時間であるため、この指標を用いて各輸送計画法の計算負荷を概算する。図 7(a) より、輸送車間再割当法による 3 つの輸送計画法は、逐次最適挿入法(処理時間:54.7 秒)と比べ、輸送計画法実行に計算時間を要していることが分かる。これは、「輸送車間に割当の範囲を拡張」したことに伴うものである。全輸送車を再割当対象とする ex2 の処理時間は、再割当対象が全輸送車 (5 台) のため、処理時間が最大 (345.5 秒) となっている。それに対し、再割当対象とする輸送車を 3 台に絞り込んでいる ex3, 4 は、割当範囲拡張に伴う計算負荷への影響は限定的 (ex3:168.9 秒、ex4:146.5 秒) である。ただし、この処理時間は最大である ex2 でも、新規デマンド 1 件当たりの輸送計画に要する時間が 2.3 秒 (345.5 秒/150) 程度である。リアルタイム性を要求しない適用シーンを踏まえると、この程度のシステム規模であれば輸送車間への割当範囲の拡張による輸送計画の計算負荷は問題ではない。しかし、ex2 と ex3,4 との計算時間の比較から類推されるように多くの輸送車からなるシステムでは、再割当対象の輸送車の台数を絞り込むことが有効になる。

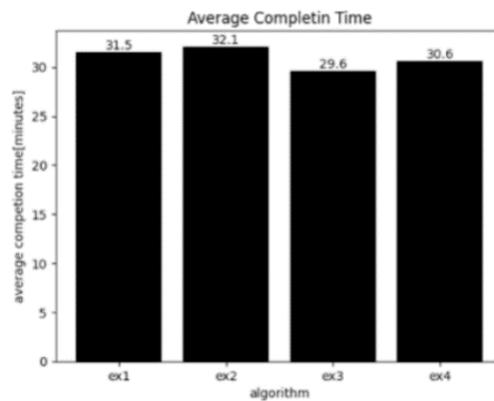
次に、図 7(b) によるデマンド達成率より、輸送車間再割当法がデマンドをキャンセルする割合を低下させる効果 (ex1:74.0% → ex2:76.7%, ex3:74.0%, ex4:75.3%) があつたことが分かる (ex3 のデマンド達成率は、ex1 と同じ)。また、図 7(c) による平均デマンド完了時間より、輸送車間再割当法の内、ex3,4 については ex1 より平均デマンド完了時間が短い。ex2 については、ex1 より平均デマンド完了時間が長い。これは、ex2 では広い範囲に拡張して良い割当を探し高いデマンド達成率を実現したため、輸送に時間を要す



(a) 処理時間



(b) デマンド達成率



(c) 平均デマンド完了時間

図7. 数値シミュレーションの評価結果

るデマンドがキャンセルされなかったことによるものである。これより、提案した輸送車両再割当法は、従来法である逐次最適挿入法より輸送計画の効率性を向上させることができることが分かった。

デマンド達成率の観点から、再割当対象の輸送車の絞り込みを行った場合は、全輸送車選定の場合程の効果は得られなかった。しかしながら、デマンド達成率・平均デマンド完了時間の両観点から、ある程度、輸送計画を効率化することができた。再割当対象の輸送車の絞り込み法は、デマンド達成率の観点からはランダム選定 (ex4) が、平均デマンド完了時間の観点からは逐次最適挿入法選定 (ex3) が良好な性質を示している。この2つの輸送車選定法は、2つの評価指標にてどちらかが優越するものではなく、比較不可能な関係にある。但し、前に述べた通り、システムのスループットに供するデマンド達成率を重視するという考えに則れば、ランダム選定が望ましいとも捉えられる。この結果は、輸送車間再割当法 (逐次最適挿入法選定) で、前処理とする逐次最適挿入法の結果に基づき選定した割当の効果がありそうな輸送車の選択が効果的でないということを意味している。今回は、1次分析に過ぎないため結論を導くには早計であるが、輸送車のランダム選定が元の輸送計画と特性が全く異なる輸送車を再割当の対象として選び出すことで隘路に陥るのを防いでいる効果によるものと類推している。

3 その他の検討状況

前章で報告した数値シミュレーション以外に、実機評価を見据えて実施している研究状況を以下で示す。

概念検証の内容検討 AMO と見立てた複数のロボット掃除機の制御に関する、概念検証のためのデモ内容 (複数機での連携による地図作成等) を検討した。

実機動作検証 LiDAR を搭載したロボット掃除機をゲーム用コントローラで遠隔操作し、ビルのフロア内を走査しながらフロア地図を作る機能を実装し、動作確認を行った (図8参考)。また、Edge と見立てた装置 (Jetson Nano) とロボット掃除機との間の通信用プログラムを作成し、通信接続とデータ送受信が行えることを確認した。

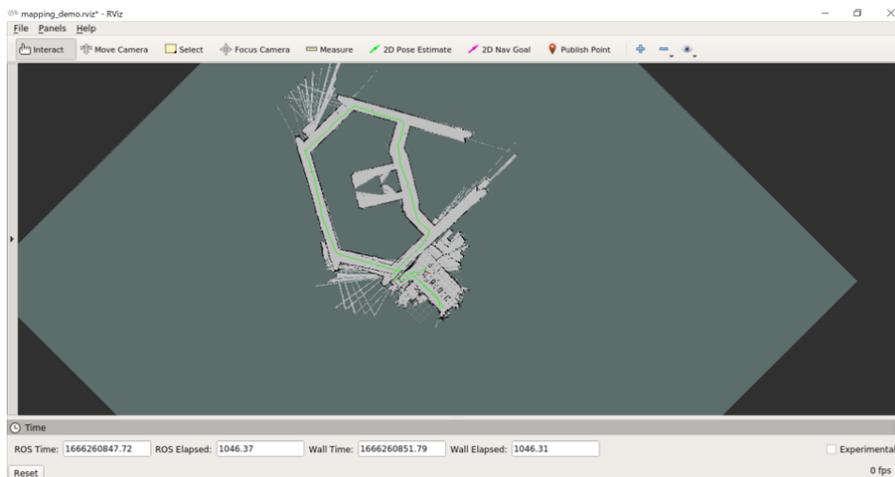


図 8. フロア地図作成の動作検証

4 おわりに

本稿では、「ヒト・モノ協調型のスマートモビリティ社会」の企画構想に基づき実施した検討の内、オンデマンド交通の輸送計画シミュレーションを中心にその検討状況を説明した。数値シミュレーションによる定量的検証と実験システムによる動作確認・概念検証の両面で、構想の有効性を確認しつつ研究を進めていく。

参考文献

- [1] 内閣府, Society 5.0
https://www8.cao.go.jp/cstp/society5_0/ (2024年3月7日閲覧)
- [2] 辻野雅之, 水上憲明, 山口直彦, 根岸豊, 二村忍, 「ヒト・モノ協調型のスマートモビリティ社会」を目指して, 東京国際工科専門職大学紀要, 第2巻第1号, pp. 44-55, 2023.
- [3] SUMO, Simulation of Urban MObility
<https://eclipse.dev/sumo/> (2024年3月7日閲覧)
- [4] 日経BP ガバメントテクノロジー, オンデマンド交通サービス
<https://project.nikkeibp.co.jp/jpgciof/atcl/19/00003/00014/> (2024年3月7日閲覧)
- [5] 野田五十樹, 篠田孝祐, 太田正幸, 中島秀之, シミュレーションによるデマンドバス利便性の評価, 情報処理学会論文誌, Vol.49, No. 1, pp. 242-252, 2008.

辻野雅之	東京国際工科専門職大学	工科学部	教授
内埜大斗	東京国際工科専門職大学	工科学部	情報工学科 4年
水上憲明	東京国際工科専門職大学	工科学部	情報工学科 講師
山口直彦	東京国際工科専門職大学	工科学部	情報工学科 助手

【研究ノート】

Sound AR サービス「Locatone（ロケトーン）」 を用いた企業連携実習の現状と展望

小野憲史・水上恵太・安彦剛志*・青山龍*・八木泉*

A prospect for corporate collaborative training using the Sound AR service “Locatone”

Kenji Ono, Keita Mizukami, Takeshi Abiko*, Ryu Aoyama*, and Izumi Yagi*

Abstract: “Locatone” is a location-based sound streaming service and platform to make people enhance walking with fun developed and operated by Sony Corporation. We have been working with it since FY2020 on a collaborative training program, where students are creating tour contents for Locatone. In this report, we will give an overview of the current status of the cooperative practice and explain the future prospects.

Keywords: Locatone, internship, game design, game education, location based game

1. はじめに

東京国際工科専門職大学では2021年度より「臨地実務実習」を展開中である。これは学生が数名ずつ企業に出向き、4週間から6週間の就業体験を行うもので、2年生、3年生、4年生の全学生を対象に実施される必修科目である。本実習において学生は、それまで授業で学んださまざまな知識や技術をもとに、各実習先で就業体験を進めていく。その一環として本学の教員である第一筆者と第二筆者[a]は、ソニー株式会社（以下ソニー）の社員である第三筆者、第四筆者、第五筆者と共同で、同社が展開するSound ARサービス「Locatone（ロケトーン）」（以下、ロケトーン）[1]の学生向けコンテンツ制作実習に取り組んできた。今回、2年生から4年生まで合計4回の実習が終了したことを受けて、これまでの成果と今後の課題について報告する。

2. ロケトーンについて

ロケトーンはソニーが開発したSound AR（現実世界に仮想世界の音が混ざり合う新感覚の音響体験）を楽しむためのサービスである（図1）。体験者は専用アプリがインストー

* ソニー株式会社

ルされたスマートフォンを片手に、ツアーと呼ばれるコンテンツを選択し、イヤホンを装着して観光地などを散策する。この時、体験者の位置情報はスマートフォンに内蔵されたGPS機能を介して、常時モニタリングされている。そして体験者が特定のスポットに到達すると、あらかじめ登録された音響情報がサーバからストリーミング再生される。これにより体験者は、音響情報という付加価値のついた散策を楽しめる。その際、音響情報の再生場所・内容・順序などを工夫することで、観光名所の情報を提供したり、ストーリーを楽しんだり、さまざまな体験を体験者に提供できる。

他にスマートフォンのカメラ機能を用いて、事前に用意された画像を実際の風景に重ね合わせて撮影する「ARカメラ」機能や、スマートフォンに内蔵されたセンサーを使って、スマートフォンを振るなど、体験者の動きをセンシングして特定の効果音を鳴らす機能、スポットごとに発行されたQRコードを読み込んで特定のイベントを再生する機能など、多彩な機能が存在する。ロケトーンは2020年11月にサービスが開始され、配信が終了したのも含めれば、ツアーの数は2024年1月現在で全国200ツアー以上にのぼる。ツアーの多くは無料で配信されているが、有料で配信されているものもある。また、博物館の音声ガイドのように、屋内施設向けのツアーとして配信されているものもある。



図1：ロケトーン（公式サイト）

3. 臨地実務実習におけるロケトーン実習について

本学では2年生を対象とした「臨地実務実習1」を1月～2月の4週間（150時間）、3年生を対象とした「2」を10月～11月の6週間（225時間）、4年生を対象とした「3」を6月～12月のうち、任意の6週間（225時間）にわたって実施している。その内容は学生が実際に企業に出向き、フルタイムで就業体験を行うものから、リモートワークや学内で実施されるものなど、多彩である。ロケトーン実習においては、第三筆者、第四筆者、第五筆者の指導のもと、学生はソニー社内で週1回のミーティングに参加し、課題説明、企画会議、ツールの使い方などの講習を受ける。それ以外の日は自宅で個人制作を行い、制作中のツアーの現地テストなどを自主的に進めていく [b]。本学教員である第一筆者と第

二筆者はミーティングに同行するほか、必要に応じて制作支援やアドバイスなどを行う。ただし、教員の指導は最低限に留め、学生の自主性が重視されている。

なお、本実習は当初、東京校のみでスタートしたが、2022年度からは名古屋校と共同で実施されている。ミーティングでは東京校の学生はソニー社内で、名古屋校の学生はオンラインで指導を受ける形で進められている。

4. 本実習の進め方について

本実習でソニー側から提示された課題は「学生一人ひとりがプロデューサーとなり、ロケトーンの可能性を広げるようなユニークなツアーを企画制作して、実際に配信すること」となる。これにより学生は実習を通して、ポートフォリオに掲載できる作品を制作し、ロケトーンの公式アプリ内で、個人名と共に配信することができる。ツアーの制作は専用のオーサリングツール「ロケトーンスタジオ」を使用して行う。ロケトーンスタジオはブラウザベースのツールで、マウスによる操作のみで、プログラミング不要で制作することができる。

実習のカリキュラムは大きく「①企画立案」「②シナリオ制作」「③素材制作」「④実装」「⑤テスト」の五段階で進められる。

このうち①と②はツアーの完成度に大きく影響を及ぼすため、全体の1/3～1/2程度の期間が費やされる。ロケトーンの特徴として、体験者が実際に街歩きをしながら、スポットごとに再生される音楽や音声を体験する点があるため、内容と場所の整合性が重要になる。その際、スポットを順番に巡らせてもいいし、順不同で巡らせてもいい。次のスポットの場所を明示せず、ヒントだけを提示して、正しい場所に体験者が移動すると、次のスポットの内容が再生される、などのツアーも制作できる。

③では音響データや、ツアー紹介のためのサムネイル画像、アイコン画像などの制作が行われる。このうち音響データは大きく「ボイス（台詞やナレーション）」「BGM」「効果音」に分類できる。ボイス制作ではフリーの合成音声エンジンによる制作に加えて、株式会社EARLY WINGの協力で、プロの声優を起用した音声収録も実施されている[c]。他に東芝デジタルソリューションズ株式会社の協力で、同社の音声合成エンジンによるボイス制作が行われたタイトルもある[d]。また、学生によっては「自分で台詞を喋って収録する」「声優志望の友人に頼む」「個人で費用を負担し、声優に依頼する」などの例も見られる。BGMや効果音、画像ファイルにおいても同様で、フリー素材などを活用するほか、イラストの描ける友人に頼む、プロのイラストレーターに発注する、などの例も見られる。

④ではロケトーンスタジオ上で制作した素材を実装していく。ツアーはロケトーンスタジオ上でプレビューできるほか、事前に登録されたスマートフォン上で実地検証ができる。ただし、スタジオ上でのプレビューと実地検証では体験が異なる。そのため学生は何度も現地に足を運び、テストを繰り返しながら、ツアーの完成度を高めていくことになる。

実習の最終日にはソニー社内で学生の最終発表が行われた後に、現地で試遊会が実施される。その後、ソニー側での最終チェックを経て、ツアーが一般公開される。ツアーは東京校の「臨地実務実習1」と「2」では新宿中央公園、「3」では北新宿4丁目商友会・親交会を舞台に制作されている。これに対して名古屋校では名古屋市内の公園や商店街むけ

に制作されている。

5. 配信済みのツアーについて

2024年1月現在で、東京と名古屋で28作品が無料配信されており、それぞれ大学のチャンネルページに一覧としてまとめられている（図2、表1～4）。このうちデジタルエンタテインメント学科の学生は22人、情報工学科の学生は6人である。ツアーは大きく物語型（複数のスポットを順番に回りながら、ストーリーを体験していくもの）と、巡回型（複数のスポットを順不同で回るもの）がある。また、ツアーの中には複数のストーリーに分岐するものや、特定のスポットを回ると、新たにスポットが出現するものなど、さまざまなバリエーションがある。「人数」はツアーがプレイされた人数（2024年1月14日現在）を示している。



図2 左からロケトーンアプリと本学の公式チャンネル、ツアーの概要ページ、ツアーの地図

表1：臨地実務実習1（2021年1月～2月）二年生

学校	ツアー名	場所	音声	概要	タイプ	学科	再生回数
東京	神隠しの爆弾	新宿中央公園	声優	公園内に仕掛けられた爆弾を歩いて解除する	物語 (分岐)	DE	93
東京	突然、中二病に話しかけられた件	新宿中央公園	声優	中二病の男性と共に公園を巡回する	物語	DE	67
東京	ナマズ	新宿中央公園	声優	猟奇殺人事件に巻き込まれる	物語	IT	70
東京	ねこさんぽ	新宿中央公園	声優	公園内の特定スポットで仮想の猫と戯れる	周遊	DE	105

表2：臨地実務実習2（2022年10月～11月）三年生

学校	ツアー名	場所	音声	概要	タイプ	学科	再生回数
東京	101枚目のデート券	新宿中央公園	声優	母親と息子の公園内での思い出を歩きながら追体験する	物語	DE	33
東京	幽霊になってもギャルでいたい	新宿中央公園	声優	霊感体質の社員が幽霊ギャルの願いを叶えていく	物語	DE	46
東京	愛の公園	新宿中央公園	声優	カップルで1つずつイヤホンを装着し、新宿の「ママ」と共に散歩デートを楽しむ	物語	IT	68
東京	えこみんを探せ！ #Locatone ゴミ拾いアクション in 新宿中央公園	新宿中央公園	声優	ゆるキャラのえこみんの雑学を聞きながら公園内のゴミ拾いを楽しむ	周遊	IT	4633

表3：臨地実務実習1（2023年1月～2月）二年生

学校	ツアー名	場所	音声	概要	タイプ	学科	再生回数
東京	トワイライトパーク	新宿中央公園	合成	異世界に迷い込んだ主人公が黒猫とともに公園を巡る	物語	DE	73
東京	Flower Talk	新宿中央公園	合成	異世界に迷い込んだ主人公が妖精とともに公園を巡る	物語	DE	10
東京	今と昔を繋ぐ新宿中央公園	新宿中央公園	合成	観光ボランティアとして老人を案内しつつ公園を巡る	物語	IT	18
東京	恋愛サーキュレーション	新宿中央公園	合成	記憶喪失となった主人公が二人の少女と公園を巡りながら、記憶を取り戻していく	物語 (分岐)	IT	61
名古屋	バグだらけのARゲームからの脱出	久屋大通公園	合成	ARゲームのテスターとなって謎を解きながらゲームのバグを直していく	物語	DE	76
名古屋	フラッと知ろう栄	久屋大通公園	合成	幼なじみの少女と共に栄地区を観光する	物語	DE	56
名古屋	コスプレイヤーとデートしたい！	久屋大通公園	声優	コスプレイヤーの少女と公園デートを楽しむ	物語	DE	35
名古屋	テレパシーハント	久屋大通公園	合成	超能力者となって周囲の心の声を聞きながら目的地をめざす	物語	DE	58
名古屋	改造少女の休日	久屋大通公園	合成	異世界から来た少女と共に公園を巡りながら休日を楽しむ	物語	DE	34

表 4：臨地実務実習 2（2023 年 11 月～12 月）三年生

学校	ツアー名	場所	音声	概要	タイプ	学科	再生回数
東京	麒麟のお世話役	新宿中央公園	声優	伝説の聖獣・麒麟の世話役となって公園を巡る	周遊	DE	5
東京	五感共有デバイステストツアー	新宿中央公園	合成	第三者に五感を共有できるデバイスの被験者として公園を巡る	物語	DE	1
東京	友のためにできること	新宿中央公園	合成	感情を失った人工知能と共に公園を巡り、感情を復活させていく	物語	DE	0
東京	国際交流パーク	新宿中央公園	声優	幽霊のカップルの雑談を聞きながら中国語を学ぶ	周遊	IT	4
名古屋	やすきち みつきちと巡る円頓寺商店街	円頓寺商店街	合成	やすきち、みつきちに案内されながら円頓寺商店街を観光する	物語 (分岐)	DE	17
名古屋	おおかみ伝承～安らぎの地へ～	円頓寺商店街	合成	呪われた少年と共に円頓寺商店街を歩き、少年の呪いを解く	物語	DE	7
名古屋	オタク君と学、円頓寺商店街	円頓寺商店街	合成	商店街オタクの少年に案内されながら円頓寺商店街を巡る	物語	DE	2

表 5：臨地実務実習 3（2023 年 11 月～2024 年 1 月）四年生

学校	ツアー名	場所	音声	概要	タイプ	学科	再生回数
東京	過去の君から	北新宿 4 丁目商友会・親交会	声優	過去にタイムトリップしたヒロインを助けるために謎解きに挑戦する	物語	DE	13
東京	猫神様へお力添え	北新宿 4 丁目商友会・親交会	声優	猫神様と共に地域猫の保護活動を疑似体験する	物語	DE	16
東京	Groovy Strolls～街と散歩とよもやま話と～	北新宿 4 丁目商友会・親交会	声優	友人と共に街歩きをしながら、地域の人々と交流していく	物語	DE	16
東京	Live or Zombie～商店街での生還記録～	北新宿 4 丁目商友会・親交会	声優	ゾンビが徘徊する商店街を突破する	物語	DE	7

6. 考察について

アーケードゲームではアクションゲームやシューティングゲームなどのように短時間で終了するコンテンツが好まれ、家庭用ゲームではロールプレイングゲームをはじめ、じっくりと遊び込めるコンテンツが適している。スマートフォン向けのソーシャルゲームでは、短時間でプレイできる一方で、シナリオやキャラクターの追加配信が可能な構造が求めら

れる。このようにデバイスやプラットフォームとコンテンツは相互に関連性を持つことが知られている。

一方でロケトーンに適したコンテンツのあり方とは何かという点について、まだ明確な回答は出ておらず、さまざまな発明や発見が続いている。こうした中、本学の学生からもいくつかユニークなアイデアが登場している。

第一に『えこみんを探せ！ #Locatone ゴミ拾いアクション in 新宿中央公園』にみられる社会課題解決型ツアーである。本作は新宿中央公園を取り囲む公道のゴミ拾いを楽しみながら行えるようにするツアーで、新宿区が主催する「第16回新宿エコワン・グランプリ」でグループ部門優秀賞に輝いた。制作した学生のSNS等での精力的な拡散活動も手伝って、試遊回数が4633回と非常に多くの体験者に親しまれている。

第二に『愛の公園』に見られるような、複数のプレイヤーで体験を共有するツアーである。本作はカップル向けのデート支援ツアーで、男女で一組のイヤホンに分け合って装着し、公園を散策する。この時、それぞれのイヤホンで異なる内容のボイスが再生され、男女の会話が自然と誘発されるような仕組みがほどこされている。今後も親子で楽しむツアーなど、さまざまな可能性が考えられる。

第三に『国際交流パーク』に見られるような、学習系ツアーである。本作は公園の各スポットを巡りながら、日本人の幽霊と台湾人の幽霊が雑談するさまを楽しむという内容である。それぞれの会話は語学学習番組のスキットを彷彿とさせるもので、楽しみながら中国語の単語や活用シーンなどが学べる（台湾人の幽霊は声優志望の留学生がボイスを担当している）。他にもさまざまな言語展開が考えられるだろう。

なお本実習は当初、デジタルエンタテインメント学科の学生を主対象としていた。しかし、これまでの実施経験では、情報工学科の学生でも特に大きなハンディは感じられなかった。これは情報工学科の学生でもゲームや映画などのエンタテインメントに関心の高い学生が志望したことが背景にあると考えられる。

7. 展望について

以上のように東京校の実習では、新宿中央公園むけにツアーが制作されてきた。これが2023年11月から2024年1月にかけて実施された臨地実務実習3では、新たに北新宿4丁目商友会・親交会の、商店街の活性化という地域課題を解決する目的で、4作のツアーが制作された。

前述のようにロケトーンには、現地を訪れなければ楽しめないという特徴がある。だからこそ現地の歴史や文脈を生かした内容にすることで、地域の観光DX推進やシビックプライドに寄与するツアー制作が可能だと考えられる。これに対して本学では、もともと地域共創デザイン実習[e]を通して北新宿4丁目商友会・親交会とのつながりがあった。そこで今回、担当教員の支援のもと、商友会・親交会へのヒアリングや、商店の協力などを経てツアーが制作された。試遊会では商友会・親交会の方々にもプレイいただき、一定の評価が得られた。今後も継続して新規ツアーを制作・配信することで、地域の魅力向上につなげていきたいと考えている。

また、ツアー制作にともない何度も現地を訪れる過程で、学生から北新宿4丁目商友会・

親交会に対する好感度が増大した、という感想が得られた。これは作り手側からの感想だが、プレイヤー側に着目した研究では、藤田（2019）がアナログの地図・冊子・WEB とスマートフォンを用いた検証を行っている。それによると地域散策支援コンテンツは舞台となった地域に対するプレイヤーのポジティブな感情生成に効果があるとしている。また、阿部・森（2019）もスマートフォンを用いた街歩き謎解きゲームのプレイを通して、参加者の地域ロイヤルティの向上が見られたと論じている。

こうした先行研究に対して、ロケトーン実習では街歩きを支援するコンテンツを学生みずから制作し、公開できるメリットがある。コンテンツの開発負荷も低いため、作り手側、遊び手側の双方を巻き込んだ、地域社会全体でのロケトーンを通じたシビックプライドの向上に寄与できる可能性がある。

また、ロケトーンでは各スポットにひもづいた QR コードを読み込んでイベントを発生させられる。そのため上下移動を伴う屋内施設など、GPS による場所計測が困難なエリアでもツアーを制作できる。そこで学生が本学のオープンキャンパスで活用するためのツアーを制作し、来校者に提供するなどの展開も考えられるだろう。今後も臨地実務実習を通して、さまざまな可能性を学生と模索していきたい。

謝辞

ツアー制作にあたり声優収録と学生指導をいただいた株式会社 EARLY WING の植木純一様、音声合成エンジンのご提供をいただいた東芝デジタルソリューションズ株式会社の河西英城様、ツアー制作においてご協力をいただいた北新宿 4 丁目商友会・親交会の皆様、学生のシナリオとキャラクターに魂を吹き込んでいただいた声優の皆様、および本実習に関してさまざまな形でご協力いただいたすべての皆様に、この場を借りて御礼を申し上げます。また、商友会・親交会とのコーディネートには本学の駒井章治教授にご協力・ご助言をいただきました。

注釈

- [a] 本実習は当初、第二筆者が大学側の担当者として進める予定だったが、体調不良等により第一筆者が引き継いだ。その後、第二筆者の体調回復に伴い、両名で実習を進めている。
- [b] その後、臨地実務実習 1（2024 年 1 月～2 月）では 3～4 名単位でのグループワークでツアーが制作されている。
- [c] 同社は声優・ナレーターのマネジメントをはじめ声優養成所、音響音楽制作・レンタルスタジオ・美容室・メディアコンテンツ制作などを手がける声優プロダクションである。
- [d] 臨地実務実習 1（2023 年 1 月～2 月）で東京校の学生向けに同社の音声合成エンジンが提供され、『今と昔を繋ぐ新宿中央公園』で使用された。
- [e] 地域共創デザイン実習とは、地域企業や自治体と本学が共同で進める Project Based Learning（PBL）形式の連携実習である。本実習で学生は企業や自治体が抱える社会課題を提示され、その解決を目的とした企画等を立案し、発表し、評価を受ける。

参考文献

- [1] ロケトーン, <<https://www.locatone.sony.net/>>（参照 2024-01-14）
- [2] 藤田美幸, 「ハイブリッドまちあるきによる地域資源の価値創造—「にいがたクエスト」を

- 事例として」、『新潟国際情報大学経営情報学部紀要2』, pp.141-151, 2019年4月.
- [3] 阿部晋吾・森一誠, 「街歩き謎解きゲームの心理的効果—気分および地域イメージの変化と、地域ロイヤルティへの影響—」, 『梅花女子大学心理こども学部紀要』 Vol.9, pp.49-58, 2019年.

【研究ノート】

留学生の就職成功に求められる 日本語能力・人材育成の効果的アプローチ — 専門学校及び日本語学校での取り組み —

山口直彦・伊勢みゆき*

Effective Approaches to Human Resource Development and Japanese Language Proficiency Required for Successful Employment of International Students

Naohiko Yamaguchi and Miyuki Ise*

Abstract: The purpose of this study is to clarify Japanese language abilities and skills and attitudes that Japanese language schools and vocational schools (or universities, etc.) must teach foreign students in order to help them find jobs in Japan. As a result of a questionnaire survey of Japanese companies and foreign students, it was found that foreign students considered the acquisition of an advanced Japanese language qualification essential and were concerned about honorifics. On the other hand, companies did not place importance on Japanese language qualifications and were concerned about internal and external communication problems due to cultural differences.

Keywords: Japanese language abilities, foreign students, career counseling

1 はじめに

著者（山口）は2016年4月から2023年3月まで、専門学校 HAL 東京（以下 HAL 東京）においてロボット工学やプログラミングなどの専門知識を講義してきた。HAL 東京では日本人だけではなく多くの留学生を受け入れており、2022年度では全体の16%にあたる566名の留学生が在籍している。山口は留学生向けの日本語専科講師補佐などの役割を通じて在学している留学生の生活指導や就職指導にも携わっていた。また伊勢はこの日本語専科の担当である。

HAL 東京に在籍する留学生の多くは日本での就職を希望している。しかし、知識・スキルが高い留学生であっても、現実では日本企業への就職に苦戦することが多い。その理由として「①日本語能力の不足」「②異文化理解の不足」「③日本の企業風土に対する理解

* 学校法人江副学園新宿日本語学校

の欠如」の3点が考えられる。卒業後の就職を目標として教育を行う専門学校として、日本人と同じく留学生に対しても就職が実現できるような指導体制を整えねばならないが、現状では課題が残されている。

2 研究目的

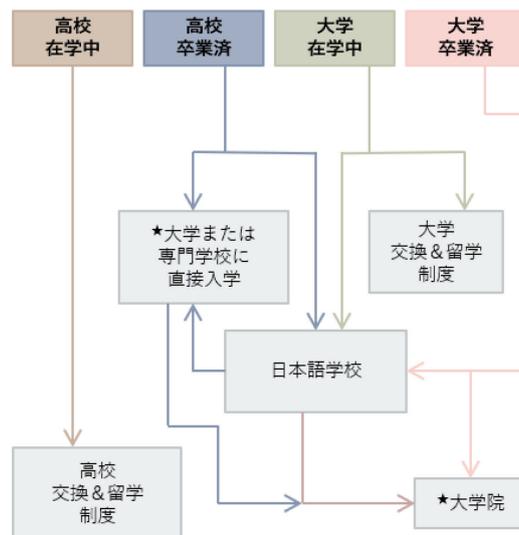
本研究の目的は、日本での就職を希望する留学生をサポートするため、日本語学校から高等教育機関（専門学校・大学・大学院。以下、専門学校等と表記）の中でどのような日本語能力やスキル及び姿勢を身につけさせるべきかを明らかにすることにある。そのため、日本の企業及び留学生にアンケートを行い、企業が求める留学生像と留学生の現状のギャップを分析することで、留学生がスムーズに日本で就職するためにどのような指導を行っていくべきかを検討する。

3 予備知識

3.1 留学のフロー

図1は海外在住の学生が日本に留学するプロセスを図示したものである。図中の交換留学制度に関しては、比較的語学力がない状態でも受け入れられ、母国に帰ることが前提となっていることが一般的であるため、本論文の検討から除外する。

日本に留学しようとする学生の多くは、多少なりとも母国で日本語の学習を行ったうえで来日する。専門学校等で学習をするために十分な日本語能力が身につけている場合は直接入学することも可能であるが、一般的には一度日本語学校に入学して、半年～2年ほど



*日本語の語学能力が求められると思われます

図1 海外在住学生の日本留学フロー（出典 [1] 山口訳）

日本語の学習を行ってから専門学校等に入学することが多い。

通常留学生は「留学ビザ」で滞在しているため、教育機関を卒業・修了した後は速やかに帰国する必要がある。配偶者ビザや永住権を取得していない留学生が学校卒業後も日本に滞在するためには、在学中に次の進学先または就職先を決めて、新しい留学ビザまたは就労ビザに切り替えなければならない。ただし、卒業後も継続して就職活動を行う者については特定活動ビザが6ヵ月単位で発行され、最長で2年間まで在留資格を得ることができる（この場合も就労ビザではないため、国内でのアルバイトは留学ビザ同様厳しく制限されている）。

以下「留学生」といった場合は現に教育機関に在籍中または教育機関卒業後就職活動を行っている人を指し、社会人経験の有無を問わず企業で働く（働こうとする）国外出身者を「海外人材」と表記する。ただし、アンケート実施時にはこのような用語の整理がまだできておらず、「海外出身者」などわずかに異なる用語が混在していたことを付記しておく。これらの用語は集計時に「留学生」「海外人材」に統一した。

3.2 日本語能力の評価

留学生の日本語能力を評価する際は、資格試験の結果を参照することが一般的である。代表的な試験として「日本語能力試験 (JLPT)【N1～N5の5段階】 [2]」「JPT 日本語能力試験【点数制】 [3]」「日本留学試験 (EJU)【点数制】 [4]」がある。これら3つの試験と、外国語学習の国際共通指標である CEFR（外国語の学習・教授・評価のためのヨーロッパ言語共通参照枠）の対応を図2に示す。ただしそれぞれ試験の性格が異なるため厳密な比較は行えず、文献によって各試験と CEFR の対応関係は大きく異なる。あくまでも参考としてとらえるべきである。

これらの資格は、専門学校等の入学条件として指定されていることが多い。

CEFR		JLPT ※	JPT	EJU
熟練した 言語使用者	C2			
	C1			
自律した 言語使用者	B2	N1		
	B1	N2	370点 以上	200点 以上
基礎段階の 言語使用者	A2	N3		
	A1	N4 N5		

※JLPTは「話す」「書く」能力を判定しないため「読む」「聴く」の2技能のみ参照

図2 CEFR と代表的な日本語能力試験の対応

4 研究方法

4.1 研究手順

本研究は、企業及び留学生に対するアンケート調査を元に行う。企業に対するアンケートでは、現在及び将来の留学生採用状況や留学生に期待する役割・能力などについて回答してもらった。留学生に対するアンケートでは、日常会話で使用する日本語と仕事で使用する日本語の違いや、就職するにあたりどのような日本語能力やスキルが必要と考えているかなどを回答してもらった。これら2つのアンケート結果を突き合せ、企業が新卒社員として求める留学生像と、現状の学生の間横たわるギャップを炙り出し、今後行っていくべき留学生教育の在り方を検討する。

4.2 アンケート調査

企業及び留学生に対するアンケートは、Google フォームを用いたオンラインアンケートによって実施した²。アンケートの実施にあたっては、集計管理の都合上個人名やメールアドレスといった個人情報を収集する必要があるため、「東京国際工科専門職大学 人を対象とする研究の倫理審査規程」に従って事前に研究推進委員会の審査を受け、承認を得た（承認番号：R2022_005）。

企業向けのアンケートに関しては、山口と交流のある企業に回答を依頼し、全17社から回答を得た。今回のアンケートは採用活動に関するものであるため、可能な限り経営者

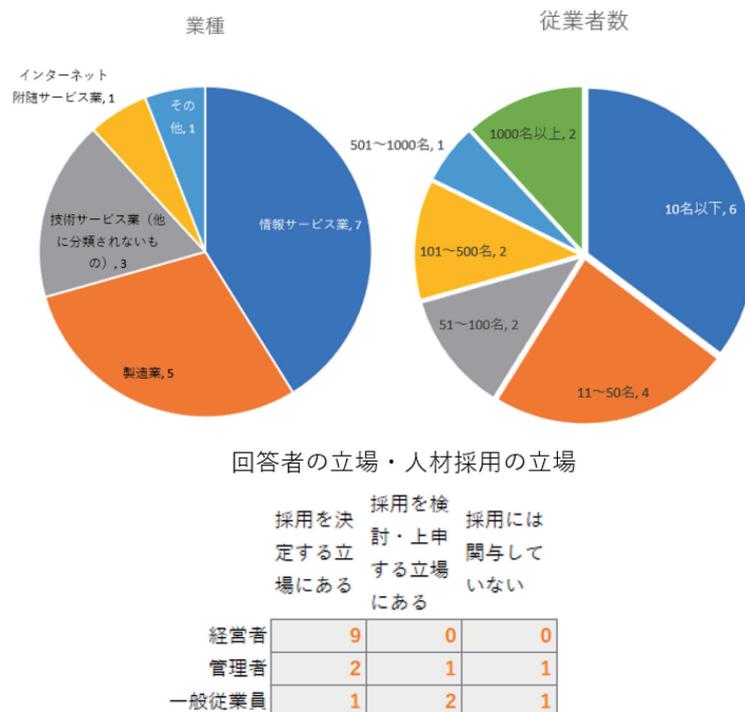


図3 アンケート回答企業概要

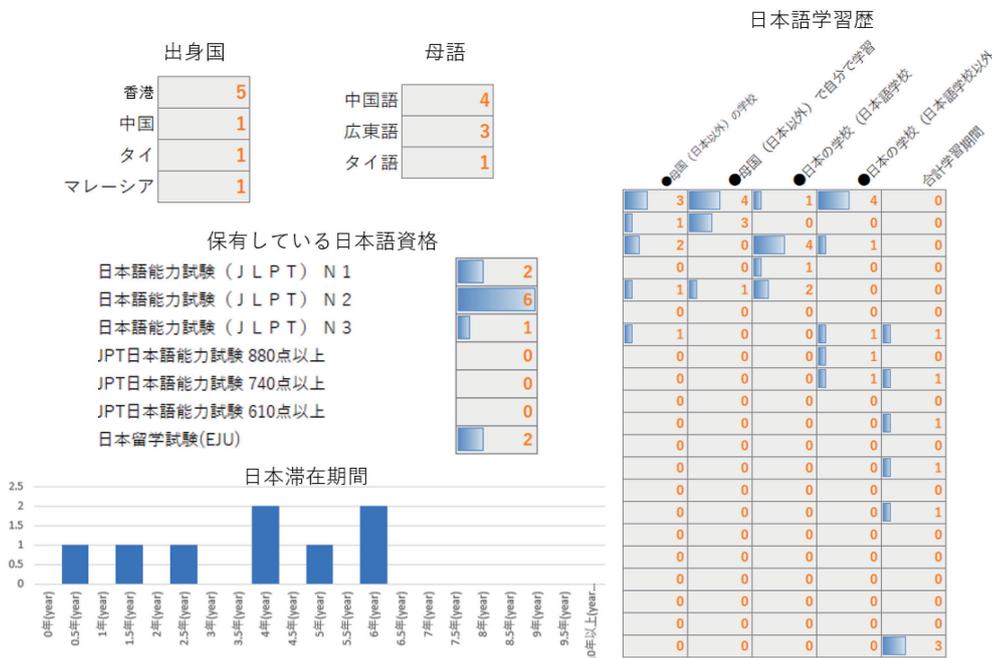


図4 アンケート回答留学生概要

または管理者として採用活動に関わる人に回答をして頂くよう依頼した。回答者が採用に直接関わらない場合は「海外人材が同僚として共に働くなら」という視点から回答をするように依頼している。アンケート回答企業の概要を図3に示す。

留学生に対するアンケートは、専門学校 HAL 東京で留学生向けに開講している日本語教育科目である「ビジネス日本語専科」「JLPT N1 対策講座」を受講している学生に回答を依頼し、計8名から回答を得た。アンケート回答学生の概要を図4に示す。

なお、本稿で示すアンケート回答の内容に関しては可能な限り原文をそのまま示している。留学生の回答に関しては日本語の誤りも含まれるが、これも原則として原文のまま示している。ただし文意を損ねないように配慮しつつ、個人の特定を避けるための修正は一部行った。

5 アンケート結果の詳細分析（企業）

5.1 企業と海外の関係性

海外との関わりが深い企業ほど、海外人材の需要があると想定されるため、分析の前情報としてまずは各企業と海外との関りについて確認しておく。

今回の調査は国内企業を対象としているため、回答を得た17社は全て本社（本店・本部）機能は日本国内に存在している。海外との関わりについて判断するため、海外支社の有無と、海外との取引（仕入れ・協力企業・顧客など）の有無について問うた。その結果を図5に示す。国内で事業が完結している企業が9社、何らかの形で海外と関係を持っている

海外取引 海外支社	海外取引なし (日本国内 で完結)	海外取引 あり
存在しない	8	0
日本国内のみ に存在する	1	4
海外に(も) 存在する	1	2
拠点によって	1	0

国内で事業が完結している	9
支社または取引で海外と関りがある	8

図5 海外支社の有無・海外取引の有無 クロス集計

企業が8社という結果になった。

また「日本語以外に業務で使用することのある言語（複数回答可）」について質問したところ、英語がもっとも多い9社、次いで中国語が4社、韓国語が2社という結果となった。

5.2 海外人材の採用状況

海外人材の採用を行っているか否かと、海外人材からの応募の有無を問い、クロス集計を行った結果を図6に示す。海外人材を採用していないが応募がある、あるいは採用しているが応募がないというミスマッチが一定程度生じていることが読み取れる。

海外人材の在籍人数については、0名が11社、1,3,4,6名が各1社であった。海外人材の出身国は中国が3名、台湾・インドネシアが各2名、韓国・マレーシア・ベトナムが各1名であった。

今後、海外人材の採用数はどのように変化するかを問うたところ、「今後増えていく」

採用状況 応募状況	海外人材の 応募がある	海外人材の 応募はない	不明
海外人材は採用していない (採用を断っている)	3	4	0
海外人材を採用している (受け入れている)	4	3	1
不明	0	1	1

：採用状況と応募状況のミスマッチ

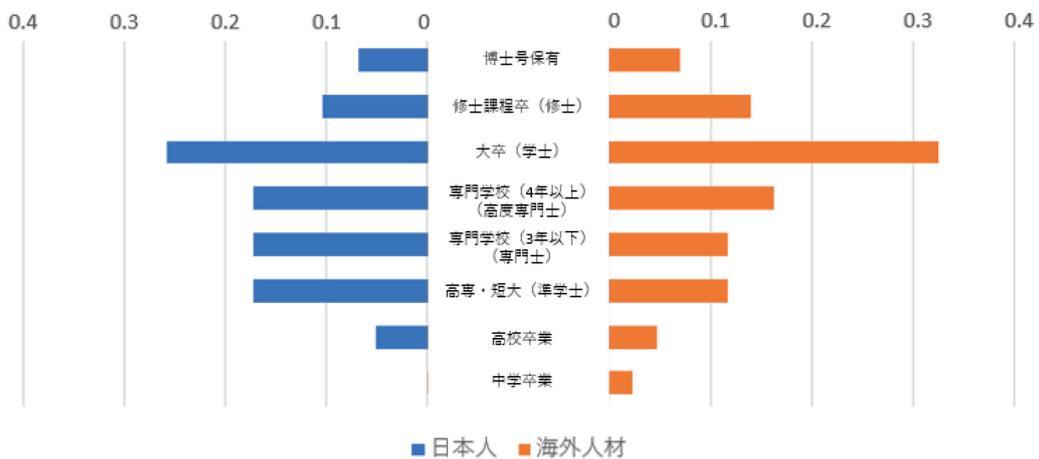
図6 海外人材の採用・応募状況 クロス集計

という回答が8社で最も多く、「特に変化しない」が6社、「今後減っていく」が1社、無回答が2社という結果となり、海外人材を現在採用していない企業であっても今後は海外人材の採用は増えていくと考える傾向が見えた。

5.3 採用者に要求する学歴とスキル

採用者に要求するスキルについて、「日本人と海外人材で要求スキルは変わらない」と回答した企業は11社、「海外人材に要求するスキルは日本人より高い」企業が4社、「海外人材に要求するスキルは日本人より低い」企業が2社であった。

採用者に要求する学歴について問うてみると、図7に示すように海外人材に対し大卒(学士)を求める企業がやや多いのと同時に、高卒・中卒を認める企業も存在することがわかる。



※日本人・海外人材各回答数の合計を母数とした比率

図7 日本人・海外人材に要求する学歴

	学位要求水準指標		
	日本人	海外人材	差異
海外人材に比べ 日本人の方が要求学歴が高い	4	2	-2
	4.166667	3.5	-0.666667
	5	4.4	-0.6
日本人と海外人材で 要求学歴は変わらない	3	2.5	-0.5
	4	4	0
	3.333333	3.333333	0
	4	4	0
日本人に比べ 海外人材の方が要求学歴が高い	3.5	3.5	0
	3.8	3.8	0
	4.333333	4.5	0.166667
	3.5	4	0.5
	4.5	5	0.5
	3.2	4	0.8
	3.2	4	0.8
	4.166667	5	0.833333
	3.666667	4.5	0.833333

要求学位に下記重みを付けて平均値を算出し、学位要求水準指標として日本人と海外人材それぞれを算出した。

- 1 中学卒業
- 2 高校卒業
- 3 高専・短大(準学士)
- 3 専門学校(3年以下)(専門士)
- 4 専門学校(4年以上)(高度専門士)
- 4 大卒(学士)
- 5 大学院修士課程・博士前期課程修了(修士)
- 6 博士号保有

図8 企業別 学位要求水準指標

各社ごとの傾向を分析するため、図8に示すような「学位要求水準指標」を定義して企業ごとの傾向を比較したところ、日本人・海外人材に対する要求学位の考え方は企業ごとに大きく異なっていることが読み取れた。海外人材に要求する日本語スキルとしては、日本語資格は必須でないという企業が14社、必須としている企業が3社という結果であり、日本語資格はそれほど重要視されていないという結果になった。日本語資格が必須であると回答した3社に要求する日本語資格を問うたところ、「日本語能力試験（JLPT）N1」が3社、「JPT 日本語能力試験 880点以上」が1社であった。

日本語能力のレベルとしては「ビジネス会話（社外との打ち合わせ・電話対応など）程度の日本語能力を必須とする」とした企業が最も多い10社、「日常会話（雑談・社内コミュニケーション）程度の日本語能力を必須とする」とした企業が6社、「高度な日本語能力（契約書の理解・翻訳、法令文書の理解・翻訳など）を必須とする」とした企業が1社であった。「日本語能力の有無は特に問わない（英語で意思疎通ができればよい、など）」と回答した企業はなかった。

その他海外人材に求める能力として自由記述をしてもらったところ、「自国の文化をきちんと日本人にわかるように説明できるスキル「論理的思考・正確さ・仕事のスピード」「日本文化の理解」という回答が得られた。

5.4 海外人材採用のメリットとデメリット

海外人材採用のメリット・デメリットのバランスについて問うた結果を図9に示す。3(中間)を中心とした正規分布様の分布となった。

「海外人材を採用することで得られると思う（または現に得られている）メリット」を問うたところ、主に「単純な「人手」として」「日本語以外の言語能力」「多様性の実現」の3点にメリットを感じていた。

「海外人材を採用することで、発生すると思う（または現に発生している）デメリット」としては主に「コミュニケーションコスト3」「文化の違いによるすれ違い」「退社時期/帰国/ビザの問題」が指摘された。

その他「特に海外人材に期待する業務・役割」「過去に海外人材を採用することで発生したトラブルや困りごと」についても自由記述によって回答してもらった。上記4つの質問に対する自由回答を図10に示す。



図9 海外人材を採用するメリット・デメリット比較

海外人材を採用することで、得ると思う（または現に得ている）	
メリット	デメリット
<ul style="list-style-type: none"> 人手不足の改善 日本人よりも採用しやすい、賃金が安い 海外企業との取引での担当 海外での日本と異なる文化や教育など持っている環境で日本人にとって新たな発見を与えてくれるのではないかと思います。 技術力の強化、翻訳など英語の活用 多様な考え方による職場雰囲気の改善 海外と業務を行う事となった場合の窓口的な動き 取引先に外資企業が多いため、英語能力がある方であれば取引先とのコミュニケーションがよりスムーズになる 社内の多様化、社外PR、社内人事整備、言語力UPや協調性UP、人材不足解消、技術承継、海外取引の可能性 技能実習生（ベトナム人）のサポートもしてくれる 日本語以外でのコミュニケーション能力向上、文化、考え方、仕事の進め方の違いによる良い影響 なし 半導体事業において海外出張の際に円滑なコミュニケーションが取れる 人材の多様性を獲得できる 海外の定性的な感覚がソフトウェア開発に重要である 	<p>コミュニケーションコストの増大 日本語でのコミュニケーションが取れば特にならない 日本人と外国人には、生活上での文化や習慣の違いがあるように仕事上でも違いがあるのかなと思います。小さな価値観のズレ、業務上のストレスや不満につながる可能性に繋がるのではないかと思います。日常会話での意思疎通が可能かどうか。また、ビジネス用語が伝わるか等 設計受託（客先のプロジェクト参加）時に国籍によってNGがある。技術流出から。 ビザについて 言語の違いによるコミュニケーションのズレ、現社員が文化の違いを受け入れるか、雇用手続き 退職の時期 コミュニケーションを取るのが難しい、勤務時間の厳守 コミュニケーションの上で正確に業務指示が伝わらない等 円滑なコミュニケーションを取るために時間を要する 開発者の採用を主に行なっているため、コミュニケーション（語学）は重要と捉える</p>
特に海外人材に期待する業務・役割	過去に海外人材を採用することで発生したトラブルや困りごと
<ul style="list-style-type: none"> 日本人にもとめていた役割と変わらない 事務作業など 単純作業 海外企業との取引での担当 日本人ではひらめくことがないような技術、アイデアが生み出され業務に反映されるのではないかと期待します。 技術面は日本人の業務と変わりありません 日本人と変わらない プログラマ、プロジェクトリーダー 英語能力 製造現場での加工業務、社内の活性化 海外の大学で学んだIT技術を活用したソフトウェア開発 積極性やバイタリティーによるチームへの影響 海外出張での円滑なコミュニケーション 出身国に関連するマーケット情報、各種情報の共有 出身地（現地）の方のニーズや感覚的な部分を開発に反映できる 	<ul style="list-style-type: none"> 帰国により定着しない。 社員間のコミュニケーションがうまくとれない 東日本大震災後、帰国してリモートワークとなってしまった 想定としては急な帰国 突然の退職 急に休んでしまう。退職してしまう。

図 10 海外人材を採用するメリット・デメリット 自由記述

6 アンケート結果の詳細分析（留学生）

6.1 日本での勤労経験

日本における勤労経験は「アルバイトとして働いたことがある」と回答した留学生が6名、「日本で働いたことはない」と回答した人が留学生は2名であった。正規職員として働いた経験がある留学生はいなかった。

「アルバイトとして働いた事がある」と回答した留学生に対し、どのような仕事をしてきたかについては「コンビニとレストランホール」「仕分け」「キッチン、接客」「コンビニエンスストアアルバイト」「セブンイレブン」「タイ語講師」の回答であった。

仕事で使用していた言語は日本語が6名、タイ語が1名で、圧倒的に日本語が多い。仕事で使用する日本語の難しさについて問う質問には「仕事で使う日本語は日常の言葉より難しい」が3名、「仕事以外の日常で使用する日本語と同じ」が2名、「日本語を使わなかった」が1名という結果であった。

日常で使用する日本語と仕事で使用する日本語の違いについては、敬語の使いかたを挙げた人が4名、業務で使用する専門用語が難しいとした人が1名であった。

言葉や文化の違いによって発生したトラブルの事例については「尊敬語正しいが使う（原

文ママ)」というものがあつた。これは意味が不明であるが、尊敬語の使いかたでトラブルが生じたものと思われる。もう一つの意見として「外国人にも伝わるように話し方が日本語より強めになります。「やすみたいです」「どうしました」より「理由は？」が伝わりやすかったです、他の人から見るとけんかになると思うことがあります。(原文ママ)」というものがあつた。これは回答者が休みを申請した側なのか受け取った側なのか不明であるが、留学生にとって分かりやすい言葉の選び方と、日本人の言葉の選び方が異なることを示す興味深い指摘である。

6.2 将来の仕事について

「将来どのような職業に就きたいと思っていますか」という質問に対しては「ゲームプログラマー」「アニメーション」「CGに関する仕事」「プログラマー」「デザイナー」「アニメーター」「研究者、ゲームプログラマー」という回答を得た。「学校卒業後、日本に長期滞在(2年以上)したいと思いませんか?」という質問には8名全員が「はい」と回答し、「将来、日本で働きたいと思いませんか?」という質問にも「はい」が7名、「どちらかといえばはい」が1名という結果となり、日本での就職を希望していることがわかる。ただし「将来、日本に永住したい(永住権を取得したい)と思いませんか?」という質問については、「はい」「どちらかといえばはい」が各3名、「どちらかといえばいいえ」が2名とややバラつく結果となった。

「日本で働きたい(または働きたくない)理由を教えてください。」という質問については、働きたい理由として「給料は本国より高い」「母国に帰りたくない」「身についたデザインの知識を活かすのが母国より日本の方が合うと思いました。」「日本のゲーム業界の発展がいいと思います」「やりたい仕事ができる」、働きたくない理由として「よく日本で働ければ、仕事の時間が長いし、休みほぼないと聞いた」との意見があつた。

6.3 海外人材が日本で働くためのスキル

「日本の会社に留学生が就職する時、必要とされる学歴はどのくらいだと思いますか?」という質問を留学生に行い、日本人に要求する学歴と比較する形で結果を図7に示した。高校又は専門学校卒であれば就職できると考えている留学生が多いことがわかる。その一方で「日本人と海外人材の採用基準(技術・技能スキル)の比較」について聞くと「海外人材に要求されるスキルは日本人より高い」と考える人が4名、

「日本人と海外人材で要求スキルは変わらない」と考える人が3名、「海外人材に要求するスキルは日本人より低い」と考える人が1名であつて、日本人に比べると高度な技術スキルを持っていないと採用されないと考えている傾向が見えた。

「日本で働くために、日本語の資格は必要だと思いますか?」という質問には「はい」が6人、「どちらかといえばはい」が2名という結果になり、全員が日本語資格の有無を重視していることがわかつた。そこで求められる(と思う)日本語資格としては図11に示す通り、日本語能力試験(JLPT)のN1またはN2が重視されていた。

「日本の会社が海外人材に求める日本語能力のレベルはどれくらいだと思いますか」という質問に対しては「ビジネス会話(社外との打ち合わせ・電話応対など)程度の日本語能力を必須とする」が4名、「日常会話(雑談・社内コミュニケーション)程度の日本語

日本語能力試験（JLPT）N1	4
日本語能力試験（JLPT）N2	3
日本語能力試験（JLPT）N3	1
JPT日本語能力試験 880点以上	1
JPT日本語能力試験 740点以上	1
JPT日本語能力試験 610点以上	0
日本留学試験(EJU)	1

図 11 日本で働くために持っていた方が良いと思う日本語資格

能力を必須とする」が3名、「高度な日本語能力（契約書の理解・翻訳、法令文書の理解・翻訳など）を必須とする」が1名という結果であり、企業側の意見と同じような分布となった。

「普段の生活（プライベートの会話や買い物など）で使う日本語と、仕事で使う日本語は違うと思いますか」という質問には「はい」が6名、「どちらかといえばはい」「どちらかといえばいいえ」が各1名となり、具体的な違いについては「ほどと（引用者注：ほとんど）同じ」「仕事の時の敬語はもっと難しいです」「敬語」「マナーが一倍（引用者注：「人一倍」のことと思われる）気にする必要があります。」「仕事で敬語を使いますが普段の生活は使わないと思います。」「仕事で使う日本語は日常で使う日本語より敬語が多い」と、敬語について指摘する意見が中心であった。

6.4 日本企業への就職

「留学生が日本企業に就職する事の難易度」について問うたところ、図 12 に示すように日本企業への就職は難しいと考えられていることがわかる。「日本人と海外人材の採用基準（技術・技能スキル）の比較」については、図 13 の通り海外人材は日本人よりも高いスキルを求められていると考えている傾向がみられた。

留学生の日本企業就職に関して自由意見を書いてもらったところ「日本文化や人間関係が合わないと思えば職場に入るともっと難しくなりますので、日本での就職をしたければ、慣れが必要です。」「日本の職場文化が厳しいと聞いていますがそうではなかったらいいと思います。」「日本人と差別がない様にしてほしい」という意見が出た。

簡単	1	1
	2	0
	3	0
	4	1
	5	4
難しい	6	1

図 12 留学生が日本企業に就職する難易度

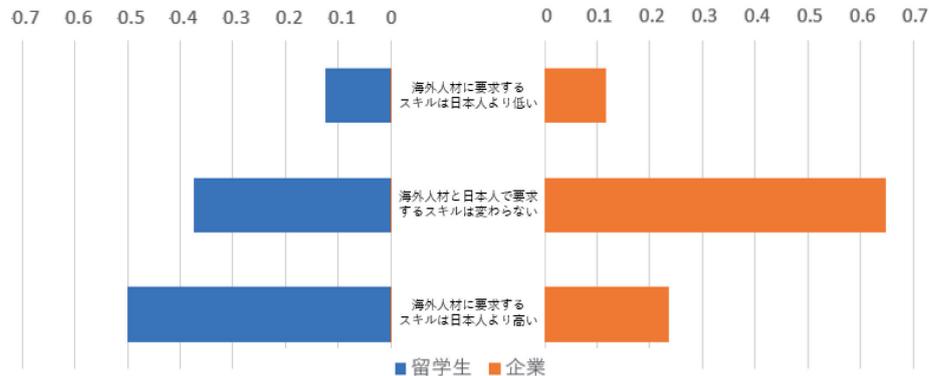


図13 日本人と海外人材の採用基準（技術・技能スキル）の比較

6.5 日本企業の社風について

日本企業と海外人材の関係性について留学生がどのように考えているのかを調べるため、図14に示す3つの質問をした。企業で働く人は多様な出身国・文化の人が集まっている方が好ましいと考えており、日本企業はもっと積極的に外国人を採用した方が良く考えている傾向が見えた。

「日本の会社に海外人材を採用して働いてもらうメリット」「日本の会社にとって、特に海外人材が活躍出来る役割や技能は何だと思いますか？⁴」「日本の会社に海外人材を採用して働いてもらうデメリット」についても自由回答してもらった。その結果を図15に示す。日本人と異なる考え方やアイデアを提供できることを海外人材が企業に提供できる大きなメリットと考えている反面、文化の衝突や交流の問題が起こりうることをデメリットとして認識していた。

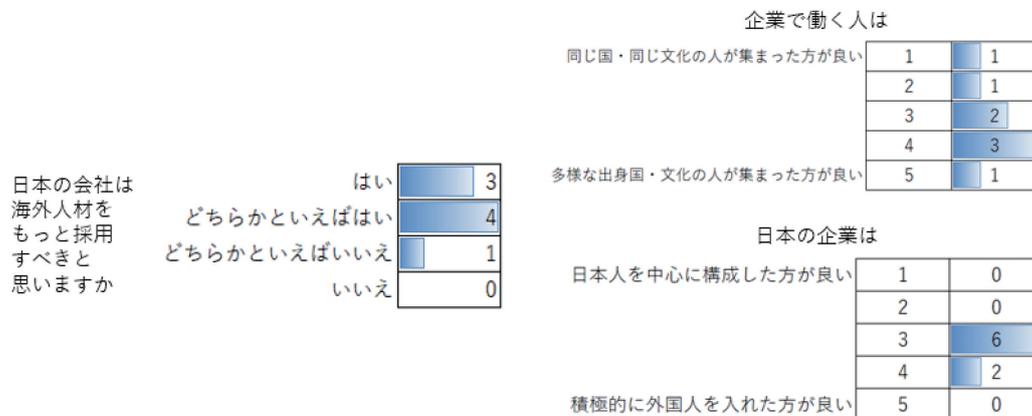


図14 留学生から見た日本企業と海外人材の関係

日本の会社に海外人材を採用して働いてもらうメリット	日本の会社に海外人材を採用して働いてもらうデメリット
<ul style="list-style-type: none"> 海外のビジネスある 多角的な考えがあります 英語能力がなくても大丈夫 グローバル展開の時に、海外の文化が知ることで視野がより幅広くなります。 国際社会との繋がりが 外国人の考え方が違うかもしれないので、ほかのいいアイデアがあるかもしれません。 日本人と違う考え方が得られる 	<ul style="list-style-type: none"> 交流の問題 なし 文化の衝突 コミュニケーションが少し難しいと思います。 文化の違いなどから誤解などを生む
日本の会社にとって、特に海外人材が活躍出来る役割や技能は何だと思いますか？	
<ul style="list-style-type: none"> コミュニケーション力、語彙力、別の立場に関する視察力 日本人と違う視点から提案できるところ：多様性、海外の文化や知識を知るきっかけになる存在、日本以外の世界の需要がわかるようになる存在 僕は外国人が日本の会社で活躍できるのは商品調整と国際貿易の部分、この国際化している21世紀、ブランドが国際化しているのかどうかはとても大事だと思う。各国の文化の異なるでお客様が気になる商品も違くなる。外国人なら、会社に自分の国の人はどういう商品が欲しいのをはっきり伝える。 私は日本の会社にとって、特に外国人が活躍できるのはコミュニケーションの能力です。会社の仕事のなかチーム製作や、他人と話し合いの場は多い、特に日本の会社ではさまざまな国の人と働くの場が多いので、よってコミュニケーション能力は日本の会社にとって重要だと思えます。 外国人は文化や成長環境が違うので、新しいアイデアを出すことは多いと思えます。会社には新しいやり方またはアイデアがあれば悪くないことと思えます。 我认为在团队亦作中需要不同性格特征的人来互相搭配，而国家不同造成的性格差异是很大的，不同的国家文化造就出不同的思维模式，这样不同的思维可似作为创新力的源泉以美国为例，他们更容易发现不同人种的特长，并把他们分到合适的岗位中。像打篮球NBA选手大多数都是运动神经发达的黑人而作为新科技发源地硅谷则大多都是黄种人、印度人。作为同一人种想要快速发展是一件很难的事，不知道不知道有没有听说过，中国的一次历史事件，闭关锁国当时的中国，就只有中国人，而因为这一政策的发布导致仅仅十年中国的科技文化生产力远远落后了别国100年。这虽然是国家级别的大事，但我觉得在会社中同样适用，我们一但适应了。这种常念，就很难有新的进步。一个好的会社环境一定不是单纯的按照原本的规则一尘不变。以我在日本打工的经验中可人看出日本人对于规则的保守。这是一件好事但他们希望外国人和他们有一样的行为准则。这时就产生了很多矛盾与问题。在我看来，有要求并不过分，是合理的，我会去接受。但我也希望对方能接受一些我的长处。比如直率的性格这在日本人里是很少有的性格特征。这个特性使我更好的“接客”受到了顾客的好评。 (訳：私はチームビルディングにおいて、異なる性格の人々がお互いに補完する必要があると考えています。異なる国の文化がもたらす性格の違いは非常に大きいです。異なる国の文化は異なる思考スタイルを生み出し、これが創造力の源泉となります。アメリカを例にとると、彼らは異なる人種の特徴をより早く見つけ出し、それを適切な職位に配置する傾向があります。例えば、NBAのバスケットボールプレイヤーは主に運動神経の発達した黒人であり、新しいテクノロジーが発祥したシリコンバレーでは主にアジア人やインド人が多く活躍しています。単一人種の社会は急速に発展するのは難しいと考えます。ご存知かもしれませんが、中国の歴史の中で一時期閉鎖国の政策が実施されました。当時の中国は中国人だけで構成され、この政策の影響でわずか10年で中国の科学技術力が他国に比べて100年も遅れる結果となりました。これは歴史の教訓であり、私は今の社会においても同じ原則が適用されると考えています。一度このような状態に慣れてしまうと、新しい一歩を踏み出すことが難しくなります。優れた社会環境は常に変化しているものです。私が日本でアルバイトの経験から見て、日本人は一般的にルールに対して保守的ですが、同じ期待を外国人にも抱くことがあります。これが多くの問題を引き起こすことがあります。私としては、ほとんどの要求とルールが合理的であると考えています。ただし、日本の方も異文化を受け入れてほしいとも期待しています。例えば、率直は日本人には少ない性格ですが、この率直な性格を持つ私は、接客時において多くの好評をいただきました。) 	

図 15 留学生からみた海外人材のメリット・デメリット

7 考察

7.1 企業は海外人材を求めるようになりつつある

アンケートの結果、現時点で海外人材を採用している企業は少ないものの、今後は採用が必要と考えている企業が多い事がわかった。その一方で、海外人材を採用していないのに応募がある、あるいは採用しているのに応募がないというミスマッチが一定程度生じていることが明らかになった事から、「海外人材と企業を適切にマッチングするしくみをさらに強化する」必要性があることがわかる。

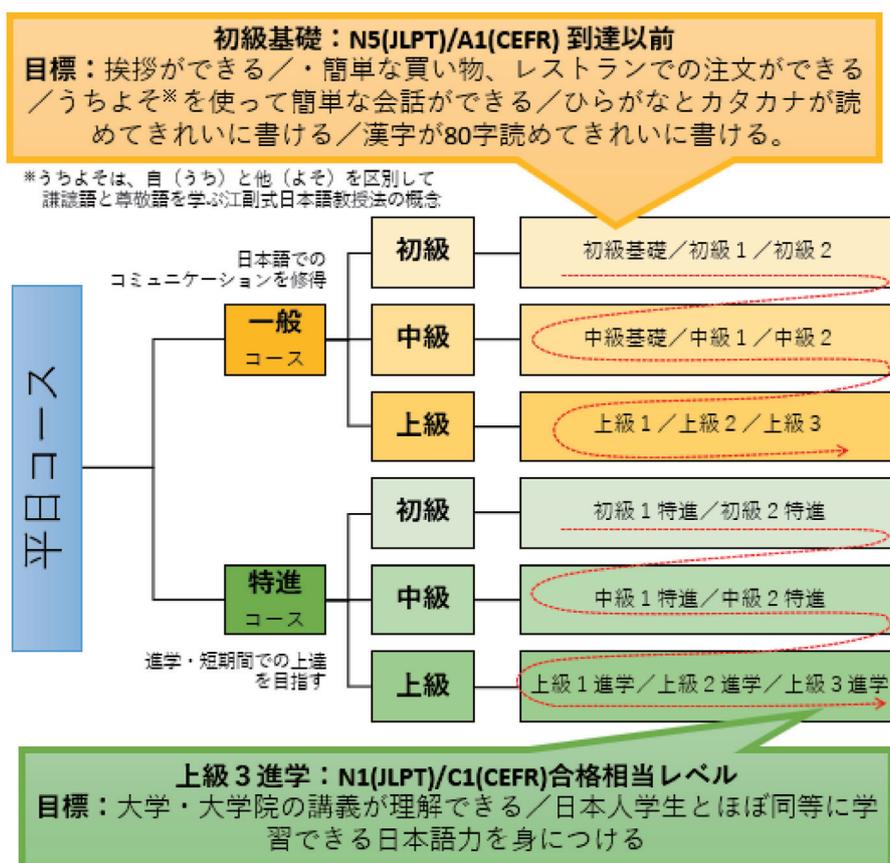
また企業側にとってはビザ関連の手続きや入社後のアフターフォローなど、受け入れたくても受け入れ態勢が（特に労務管理や人事面で）整わないという意見があった。この点に関しては教育機関からの積極的な対策は難しいものの、国や地方自治体あるいは民間の

ビジネスコンサルタント企業等を通じて「企業が海外人材を採用するサポート」を行う体制を整えていく必要があると考えられる。

7.2 敬語=正しい日本語と言う意識からの脱却指導

日本語学校に通う留学生は、通う目的もスタート時のレベルも様々である。

一例として、著者(伊勢)が勤務する新宿日本語学校のカリキュラム概要を図16に示す。学校紹介ページにおいて「新宿日本語学校(SNG)は、日本語での会話、日本の大学進学や日本での就職をお手伝いします。[6]」とある通り、日本での日常生活を送ることから、進学や就職を目的とした学習までを日本語学校ではカバーする必要があり、これを実現するためのカリキュラムが組まれている。新宿日本語学校で一番中心となるカリキュラムは平日コース(平日の昼間に学ぶ)で、さらに日本語でコミュニケーションができるようになることを目指す「一般コース」と日本の大学、大学院、専門学校への進学や短期間で日本語上達を目指す「特進コース」に分かれ、個々の能力や目標に合わせてステップアップを目指している。



学習開始時の能力に合わせたクラスから始め、ステップアップする。
 この他目的に応じて「ビジネス日本語」や「日本留学試験対策」等の
 追加コースを受講可能

図16 新宿日本語学校 平日コースカリキュラム概要図

このように、日本語学校では可能な限り学生のニーズを汲んだ教育が受けられるカリキュラムを組んでいるものの、現実問題として学生指導やスキル評価の物差しとしては、資格（日本語能力試験（JLPT）他）による能力判断が用いられる。これらの試験では「いかに正しい日本語を使えるようになるか」が重視され、その中に「正しい敬語」も含まれている。（日本語学校在学中も、専門学校等へ進学した後も）留学生には「正しい日本語＝正しい敬語」という意識が刷り込まれている。

もちろん正しい敬語が使えるに越したことはないし、専門学校等への進学に一定レベル以上の日本語資格が求められている以上、日本語資格対策の教育が重要であることは否定しない。だがそれ以前に、学校を卒業・修了した後に日本の社会人として巣立っていくためには、日本文化に合わせた日本流のコミュニケーション術をしっかりと身につけておく必要がある。これは（文法的に）正しい敬語を学ぶだけでは不十分である。

たとえ正しい敬語であっても、使う相手やシチュエーションを間違えれば失礼に当たってしまう。例えばアポイントをとって人に会うとして、「本日はお忙しい中、お時間をいただき、誠にありがとうございます。」という挨拶は敬語としては正しい。伝える相手が自社の社長や役員クラスの人、あるいは外部企業であれば全く問題がないが、直属の上司や同じ部署の先輩に使うのは場違いな発言になってしまう。一步間違えば慇懃無礼で失礼な態度と思われる危険性すらある。このような言葉選びの判断を学生が学ぶことは非常に難しいが、社会人として日本のコミュニティに馴染んでいくためには乗り越えねばならないハードルである。また高度コミュニケーションのみならず、日常的なコミュニケーションの前提として、日本文化と自国文化のギャップについてもよく理解しておく必要がある。

7.3 日本語能力試験（JLPT）にこだわりすぎない教育

日本語学校から専門学校等へ進学するにあたり、留学生は日本語能力試験（JLPT）N1等の日本語資格取得を目指して学ぶ。その影響からか、就職に際しても企業は海外人材に日本語資格を求めていると留学生が考えていることが今回のアンケート結果からわかった。また、HAL 東京では学校としても留学生に日本語能力試験（JLPT）N1を目指すように推奨しており、学生の認識を再度裏付ける（強化する）体制になっている。

しかし今回行ったアンケートでは、企業側は日本語資格の有無はさほど気にしておらず、ビジネス会話や社内でのコミュニケーションがしっかりできることを重視していた。専門学校等での日本語指導にあたっては、資格取得と同等かそれ以上に、実践的なビジネスコミュニケーションを学ばせる必要があると思われる。日本語能力試験（JLPT）に比べ知名度はまだ低い、ビジネス分野に特化した日本語資格（例：日本漢字能力検定協会「BJT ビジネス日本語能力テスト [5]」）も行われているため、学校・企業ともにこのような資格を教育や能力判断の参考にするのも有効であろう。

7.4 企業文化のすり合わせ

企業側も留学生側も「（企業の）マナーや文化・習慣を守れるか」を不安要素として挙げている。「郷に入っては郷に従え」という言葉がある通り、「〈外／後〉から来た者が〈中／先〉の者のやり方に合わせる」ことが組織や社会の中でうまく立ち回るためのセオリーではある。しかし、海外人材と日本企業の間横たわる文化や考え方のギャップがあまり

に大きい場合、双方に大きな労力を課することになる。

専門学校等での就職指導にあたっては、日本の社会・企業の文化や考え方を十分に伝え、日本社会の中で戸惑わないように導くことが必要であるが、同時に企業側も従来の（日本人の）やり方を無理に押し付けるのではなく、留学生（海外人材）がなじみやすいような会社の在り方をもう少し柔軟に考えるべきである。双方が歩み寄って〈外／中〉と〈先／後〉の間に存在するギャップを減らすことが、結局のところ、双方幸せになる道である。

8 今後の課題

8.1 今後の研究に向けて

本研究で調査は、企業17社、留学生8名という極めて小規模なアンケートであり、対象企業もIT系企業に絞って行っている。また対象分野（業種・進路）もさらに広げていく事で、分野ごとの傾向差が見える可能性がある。より普遍的な分析を行うため、企業・留学生共に対象数を増やした本格的な調査を実施する必要がある。

8.2 今後の教育に向けて

今後は今回のアンケートをもとにして留学生（海外人材）と企業双方の間のギャップを埋めることを目的とした授業に取り組んでいく必要があるだろう。そのためには段階を踏んだ就職活動を見据えた授業が必要であり、専門学校で学ぶ留学生たちには早い段階での意識づけが必要であると思われる。

本研究の成果をベースとして、具体的・実践的な留学生指導カリキュラムの検討や教材作成を試みていきたい。

9 おわりに

本研究は、企業及び留学生に対するアンケート調査を行い、企業からみた海外人材採用状況の実態と、留学生が考える就職観の調査を行った。

その結果、企業としては海外人材を採用する必要性を認識しているものの、実際に採用している企業はまだ少ない。海外人材の採用にあたっては日本の企業文化（ビジネスマナー等）になじめるか否かを懸念していた。

その一方で留学生もまた、就職にあたって日本の企業文化（ビジネスマナー等）の理解が難しいと感じていることが分かった。加えて就職に際し留学生は日本語資格が必須と考えているが、企業側はそれほど日本語資格を求めているというギャップも見えた。

留学生のアンケートでは、日本語のトラブルだけではなく、それ以前のコミュニケーションに問題があるのでは？と思われる部分も多い。留学生は「ビジネス日本語＝敬語の使用」ととらえていることが多いが、現実にはさらに高度なコミュニケーションを要求される。企業側からのアンケートにもあったように、日本の企業文化、また職場での人間関係の指導を行っていくことが、留学生を送り出す教育機関としては重要であることがわかった。

謝辞

中国語の翻訳には、HAL 東京先端ロボット開発学科卒業生の隆家琪さんに協力頂きました。

参考文献

- [1] 在アメリカ合衆国日本国大使館「STUDY IN JAPAN: A Guide for U.S. Students」、https://www.us.emb-japan.go.jp/jicc/doc/Library/StudyAbroadGuide_Spreads.pdf、p.9、2024年01月11日確認。
- [2] 「日本語能力試験 JLPT」、<https://www.jlpt.jp/>、2024年01月11日確認。
- [3] 「JPT 日本語能力試験」、<https://www.jpctest.jp/>、2024年01月11日確認。
- [4] 日本学生支援機構「日本留学試験 (EJU)」<https://www.jasso.go.jp/ryugaku/eju/index.html>、2024年01月11日確認。
- [5] 日本漢字能力検定協会「BJT ビジネス日本語能力テスト」<https://www.kanken.or.jp/bjt/>、2024年02月05日確認。
- [6] 新宿日本語学校「新宿日本語学校について」<https://www.sng.ac.jp/jp/about/>、2024年02月10日確認。

注

- 1 日本留学試験 (EJU) は日本の大学等で必要とする日本語力及び基礎学力の評価を行うための試験であって日本語能力のみを評価するものではないが、日本語科目の成績をもって日本語能力の評価指標とする場合がある。
- 2 ただし、留学生向けアンケートの一部設問に誤植があり、正しい回答が得られなかったため、当該質問だけは後日紙ベースで再回答してもらった。
- 3 ここでいう「コスト」は金銭的なものではなく、コミュニケーションを成立させるために必要な労力・時間・手間などを総称して表現しているものと想像される。
- 4 この設問のみ、オンラインアンケート時の設問不備があったため、後日紙で再回答してもらった。

【研究ノート】

生成 AI を活用した未知の領域の学習への試み

—量子コンピューティングを例にして—

廣峯優希・富山哲男・宮崎淳*

**Attempts to learn the unknown by using generative AI
: For example, Quantum Computing**

Yuki Hiromine, Tetsuo Tomiyama, and Jun Miyazaki

Abstract: This paper concentrates on the application of generative AI in learning quantum computing. It discusses guidelines, usage examples, and precautions for beginners in quantum computing by utilizing generative AI. Moreover, it presents instances of utilizing generative AI in learning quantum computing, offering objective perspectives on its potential uses in various research and educational fields.

Keywords: Quantum Computing, generative AI, QAOA

1. はじめに

2022年にOpenAIが発表したChatGPT[1]をきっかけに、世界的な生成AIブームが巻き起こっている。生成AIは、個人的な使用を目的とするユーザだけでなく、教育や研究などにも活用されている。文部科学省では、2023年の7月4日に“初等中等教育段階における生成AIの利用に関する暫定的なガイドライン”[2]を公開しており、国内でも政府が主導して生成AIを活用しようとする動きが見られる。

2. 生成AIを使って未知の領域の学習・研究にチャレンジする—量子コンピューティングを例として—

本論文では、2023年11月現在公開されている生成AIのうち、OpenAIのChatGPT[1]と、MicrosoftのBingAI[3]、GoogleのGoogleBard[4]を使用して、量子コンピューティングの学問に触れたことがない初学者を支援し、生成AIの活用方法についてのガイドラインを提供するとともに、今回の実験に使用する各々の生成AIの性能の比較について考察を行うことを目的としている。

2011年にカナダのD-Wave System社が世界初の商用量子コンピュータを販売[5]して以

* 東京国際工科専門職大学客員教授、株式会社オレンジテクラボ

降、量子コンピューティングは世界中でめざましい研究開発が行われている。さらに、量子コンピュータは量子力学の重ね合わせを利用 [6] し、一部の問題で計算量が減り、計算が高速化するマシンとして期待されている。[7] しかしながら、現在一般的に使用されている従来のコンピュータ（以降「古典コンピュータ」とよぶ）と量子コンピュータを比較すると、量子コンピュータはハードウェアの制約や研究開発段階といったこともあり、古典コンピュータを超える計算速度を誇るほどの実用化には至っていない。また、量子コンピューティングは、量子力学の理論をもとに研究されているため、量子力学の理論が完全でないことから未知の領域を含んでいることに加え、人間が直感で理解できる古典力学とは理論が大きく異なってくる。今回は、このような未知の領域の学習・研究にチャレンジし、そこで行った実験や学習の成果について考察をしていく。

3. 量子計算の基礎

量子コンピュータを使った量子計算の基礎について、一般的に使用されている古典コンピュータと比較して、2つの項目に分けて説明する。

3.1 情報の最小単位（ビット）

古典コンピュータは、電圧の高低を利用して、電圧が低いときは0、電圧が高いときは1として古典ビットを表現してきた。量子コンピュータの場合は、電子スピンの状態 [8] によって量子ビットを表現する。古典ビットと量子ビットの表現について、以下の図1に示している。

古典ビット	量子ビット
<p>【定義】 電圧の高低によって0,1の状態が表される。</p> <p style="text-align: center;">   </p> <p>低電圧"0" 高電圧"1"</p>	<p>【定義】 電子スピン(角運動量)の状態によって、重ね合わせの状態が表される。</p> <p style="text-align: center;">    </p> <p>アップスピン" 0>" ダウンスピン" 1>" 重ね合わせ" ψ>"</p>

図1. 古典ビットと量子ビットの表現

以上の図1より、古典コンピュータでの古典ビット"0","1"という表現は、量子ビットでは"|0>","|1>"に対応する。また、量子ビットでは量子力学の原理により重ね合わせの状態をとることができる。重ね合わせによって表された量子ビットの状態|ψ>は以下のように表される。

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

ここで、 α, β は複素数で $|\alpha|^2 + |\beta|^2 = 1$ である。また|ψ>は量子状態を表す複素ベクトルである。

ここでいう“原理”とは、他の基本原理から導出できないもので、それに反する実験事実もないことから、それを疑っても仕方がないという意味であるため、量子力学の原理を受け入れて理解を進める必要があることに注意しなければならない。

さらに、それぞれの量子ビット $|0\rangle, |1\rangle$ について、以下で定義される。

$$|0\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

3.2 論理回路とその出力パターン

古典コンピュータには古典回路、量子コンピュータには量子回路と呼ばれるものがある。量子計算では次のように定義されるテンソル積が使用される。

$$U \otimes V := \begin{pmatrix} u_{00}V & u_{01}V \\ u_{10}V & u_{11}V \end{pmatrix} = \begin{pmatrix} u_{00}v_{00} & u_{00}v_{01} & u_{01}v_{00} & u_{01}v_{01} \\ u_{00}v_{10} & u_{00}v_{11} & u_{01}v_{10} & u_{01}v_{11} \\ u_{10}v_{00} & u_{10}v_{01} & u_{11}v_{00} & u_{11}v_{01} \\ u_{10}v_{10} & u_{10}v_{11} & u_{11}v_{10} & u_{11}v_{11} \end{pmatrix}$$

また、テンソル積は以下の性質を満たす。

$$\begin{aligned} \alpha(U \otimes V) &= (\alpha U) \otimes V = U \otimes (\alpha V) \\ (U_0 + U_1) \otimes V &= U_0 \otimes V + U_1 \otimes V \\ U_0 \otimes (V_0 + V_1) &= U_0 \otimes V_0 + U_0 \otimes V_1 \end{aligned}$$

これらのテンソル積の定義と性質を踏まえた上で、古典回路と量子回路の違いについて以下の図2で示す。

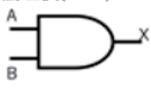
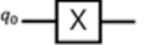
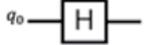
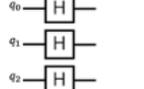
古典回路	量子回路
【回路記号】 論理積(AND)  論理否定(NOT)  ...など	【量子ゲート】 Xゲート(NOTゲート)  Hゲート(アダマールゲート)  ※ただし、 $ 0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $ 1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$...など
【論理式】 論理積(AND) $X = A \cdot B$ 論理否定(NOT) $X = \bar{A}$	【行列表現】 Xゲート(NOTゲート) $q_0 \otimes X = q_0\rangle \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ Hゲート(アダマールゲート) $q_0 \otimes H = q_0\rangle \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
・古典回路の例(3 bit) $A=0, B=0, C=0$ のとき  $X = AB + \bar{A}\bar{B} + C$ $X = 0 \cdot 0 + 1 \cdot 1 + 0$ $X = 1$ よって、出力は"X=1"	・量子回路の例(3 qubit) $q_0 = 0\rangle, q_1 = 0\rangle, q_2 = 0\rangle$  $(q_0 \otimes H) \otimes (q_1 \otimes H) \otimes (q_2 \otimes H)$ $= (0\rangle \otimes H) \otimes (0\rangle \otimes H) \otimes (0\rangle \otimes H)$ $= \frac{1}{\sqrt{8}}(000\rangle + 001\rangle + 010\rangle + 011\rangle + 100\rangle + 101\rangle + 110\rangle + 111\rangle)$ よって、量子ビットの状態は"8つの状態の重ね合わせとなる"

図2. 古典回路と量子回路の違い

以上の図2では、古典回路と量子回路における回路の表し方や数学的表現、回路の例について表している。古典回路と量子回路の決定的な違いとして、古典回路はある入力に対して出力は1つに定まるが、量子回路ではある入力に対し、量子ビットの状態は確率的表現で表されることである。

4. 量子アルゴリズム

量子計算におけるアルゴリズムの意義について、Quantum Native Dojo[9] に解説がある。

【第2章 量子アルゴリズム入門】

量子コンピュータは、量子力学的な重ね合わせによって、 n 個の量子ビットを用いて 2^n 個の状態を同時に処理できる。しかし、これだけでは「計算が速い」ということにはならない。なぜなら、計算終了後に結果を観測する際に、 2^n 個の状態の内どれか一つがランダムに得られるのみだからである。したがって、欲しい答えが高確率で得られるように設計された、量子コンピュータ専用のアルゴリズムが不可欠である。そのようなアルゴリズムを量子アルゴリズムと呼ぶ。

このことから、組み合わせ最適化問題の解を求めるためには、量子アルゴリズムを使用する必要がある。

量子計算で使用されるアルゴリズムについて、Quantum Algorithm Zoo[10] で 2022 年 6 月 26 日時点において約 60 種類ほど公開されている。[10] で公開されている量子アルゴリズムに関して、以下の 4 つの種類に分類されている。

1. 代数的整数論アルゴリズム (Algebraic & Number Theoretic)
2. オラクルアルゴリズム (Oracular)
3. 近似アルゴリズム (Approximation and Simulation)
4. 機械学習と数値最適化 (Optimization, Numerics & Machine Learning)

ここで、量子計算で使用されるアルゴリズムについて、3 に分類される QAOA (Quantum Approximation Optimization Algorithm)[11] を例に解説していく。QAOA は量子ゲート方式において、組み合わせ最適化問題を解くためのアルゴリズムの一つである。このアルゴリズムの理論を、“IBM Quantum で学ぶ量子コンピュータ [12]” から引用したものを以下に示す。

【QAOA の手順】

- (i) 初期状態 $|s\rangle = |+\rangle^{\otimes n}$ を用意する
- (ii) パラメータ $\{\beta_1, \dots, \beta_l\}$ と $\{\gamma_1, \dots, \gamma_l\}$ で指定される 2 つの入力 $U_C(\gamma_i), U_X(\beta_i)$ を作用させ状態 $|s\rangle \rightarrow |\beta, \gamma\rangle$ を得る。
- (iii) 量子コンピュータで期待値 $\langle \beta, \gamma | C(Z) | \beta, \gamma \rangle$ を評価 (測定) する。
- (iv) $\langle \beta, \gamma | C(Z) | \beta, \gamma \rangle$ を小さくするように古典コンピュータでパラメータ β, γ を更新する。
- (v) 最適解 β^*, γ^* を得るまで (i) ~ (iv) を繰り返す。
- (vi) 何回か状態 $|\beta^*, \gamma^*\rangle$ を Z 基底で測定し、良さそうな測定結果 (ビット列 z_1, \dots, z_n) を解として出力する。

QAOA のアルゴリズムの手順は以上のようにになっているが、これでは理解が難しいので、先ほどの QAOA の手順の一部を以下の図 3[13] に対応させる。QAOA の手順における (i) ~ (v) を以下の図 3 の "1 ~ 5" に対応させ、QAOA の手順 (vi) を最終的な出力結果として表す。

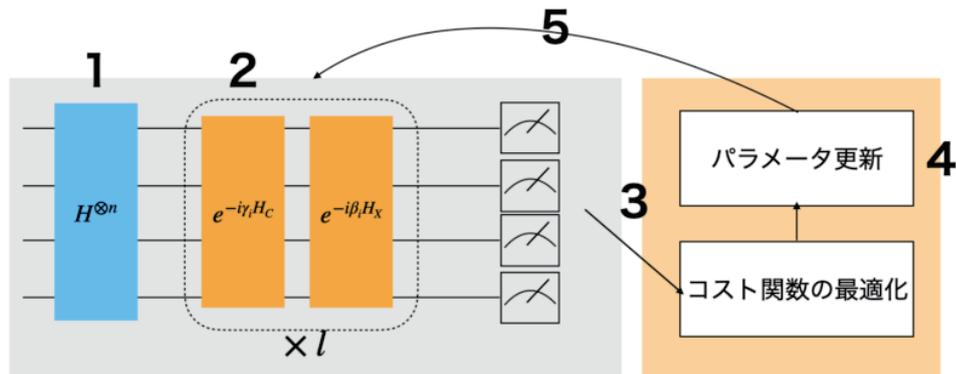


図 3.QAOA アルゴリズムの手順 [13]

この図 3[13] を用いて、量子アルゴリズムについて例題を以下で取り扱う。例題として、ここでは最大カット問題を取り扱う。最大カット問題は、重みつきグラフの辺を 2 つのグループに分割する問題で、グループ分けをした際に分割される辺の数を最大化するものである [14]。今回は、以下の図 4 ~ 図 7 を例に説明する。

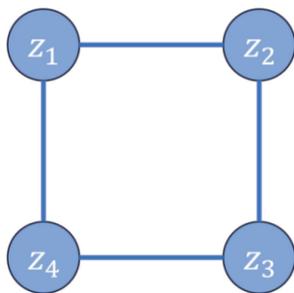


図 4. 無向グラフ

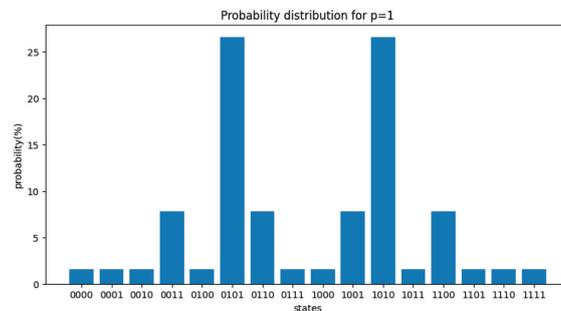


図 5. 2 つの入力 (γ_i), $U_X(\beta_i)$ を 1 度作用させたもの

この図 4 のグラフは、初期状態 (i) を表している。ここで、2 つの入力 (γ_i), $U_X(\beta_i)$ を作用 (ii) させると、図 5 のような結果が観測される (iii)。さらに、2 つの入力 (γ_i), $U_X(\beta_i)$ を作用させる回数を増やす (iv)(v) と、図 6 のように最適解を得ることができる (vi)。

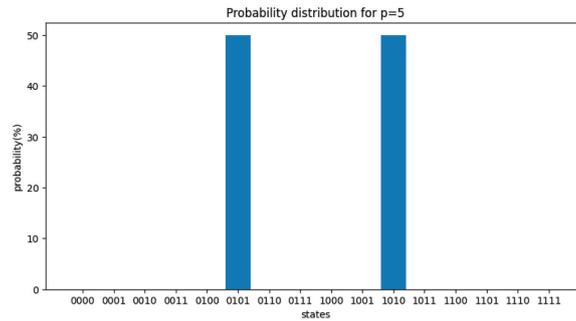


図 6. 2つの入力 $(\gamma_i), U_X(\beta_i)$ を複数回作用させたもの

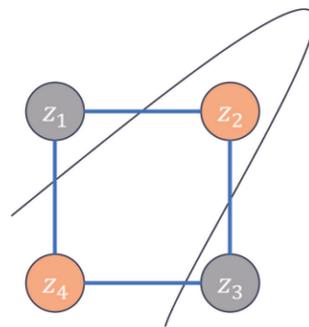


図 7. 得られた最適解

ここで図 6 と図 7 をみると、図 6 で $|0101\rangle$ と $|1010\rangle$ が得られた。この 2 つの最適解のうち $|0101\rangle$ の解を用いると、 z_1 と z_3 のグループと z_2 と z_4 のグループに分けて、図 7 のように辺を 4 つに分割すると、辺の分割を最大化することができる。

このように、量子計算特有の重ね合わせを利用することで、身近な問題の最適解を求めることができる。

5. 生成 AI の自己学習への活用

筆者の一人（廣峯）は、2023 年 10 月 2 日から同年 11 月 14 日まで「生成 AI を活用して量子コンピューティングを自己学習する」ことをテーマにインターンシップの体験をした。その中で、全 6 週間におよぶインターンシップの中で生成 AI を活用した場面は、以下のとおりである。

1. 6 週間の実習計画表の作成
2. 生成 AI を使った Python コードの生成

上記の 1,2 について、どのように生成 AI を活用してどのような結果が得られたかについて、以下で説明する。

5.1 6 週間の実習計画表の作成

まず、6 週間分の実習計画を立てるために、ChatGPT(GPT-4.0)[1] を活用して、実習計画

表の作成を行なった。本インターンシップでは以下のプロンプト (図 8) に対する出力結果 (図 9) をベースとして、実習計画書の進捗チェックリストを作成し、図 10 の実習計画表に従って実習を進めていった。



6週間のインターンシップを実施します。
その6週間分のインターン実施計画を作成してください。表や図を使っても良いです。
内容は主に量子コンピュータを扱い、その分野の中でQAOAのような量子ゲートにおける最適化の学習や、ChatGPT(GPT-3.5)やBingAI(GPT-4.0)を用いた量子ゲートにおけるプログラミングの支援に関する事柄について学習します。

図 8. 実習計画書の作成にあたって使用したプロンプト

<p>週1-2: 量子コンピューティング基礎</p> <ul style="list-style-type: none"> 日程: 週1-2 内容: <ul style="list-style-type: none"> 量子コンピューティングの基本概念の学習 量子ビットと量子ゲートの理解 Qiskitなどの量子コンピューティングフレームワークの導入 	<p>週7-8: プロジェクトの準備と設計</p> <ul style="list-style-type: none"> 日程: 週7-8 内容: <ul style="list-style-type: none"> インターンシッププロジェクトの具体的な目標の設定 必要なリソースとデータの収集 プロジェクトの計画と設計の作成
<p>週3-4: 量子ゲート最適化 (QAOA)</p> <ul style="list-style-type: none"> 日程: 週3-4 内容: <ul style="list-style-type: none"> Quantum Approximate Optimization Algorithm (QAOA) の理論と実験の学習 QAOAを使用した最適化問題の解決方法の習得 QAOAの実際の問題への適用 	<p>週9-10: プロジェクト実施と実験</p> <ul style="list-style-type: none"> 日程: 週9-10 内容: <ul style="list-style-type: none"> プロジェクトの実験と実験の開始 ChatGPTまたはBingAIを使用した量子ゲートプログラミング支援の実際の適用 問題の特定と解決
<p>週5-6: ChatGPTとBingAIの活用</p> <ul style="list-style-type: none"> 日程: 週5-6 内容: <ul style="list-style-type: none"> ChatGPT (GPT-3.5) およびBingAI (GPT-4.0) の理解と機能の調査 量子ゲートプログラミング支援用のGPTモデルのカスタマイズ方法の学習 プロジェクトにおけるChatGPTまたはBingAIの活用 	<p>週11-12: プロジェクト評価と成果報告</p> <ul style="list-style-type: none"> 日程: 週11-12 内容: <ul style="list-style-type: none"> プロジェクトの評価と結果の分析 結果をドキュメント化し、成果報告の準備 インターンシップの成果を発表するためのプレゼンテーションの準備

図 9. 図 8 のプロンプトに対する出力

<p>進捗チェックシート</p> <p>週1: 量子コンピューティング入門</p> <ul style="list-style-type: none"> Day 1-3: 量子コンピューティングの基本概念を学ぶ <ul style="list-style-type: none"> <input type="checkbox"/> 基本的な量子力学の概念の理解 <input type="checkbox"/> 量子計算モデルの基礎を把握 Day 4-7: 量子ゲートと量子ビットの操作に関する理解を深める <ul style="list-style-type: none"> <input type="checkbox"/> 主要な量子ゲートを学ぶ <input type="checkbox"/> 量子回路の構築と実験 <p>週2: QAOAの理論</p> <ul style="list-style-type: none"> Day 8-10: QAOAの理論を学ぶ <ul style="list-style-type: none"> <input type="checkbox"/> QAOAの基本的な仕組みの理解 <input type="checkbox"/> QAOAの利点と制限を把握 Day 11-14: QAOAアルゴリズムの詳細な理解と数学的背景の探求 <ul style="list-style-type: none"> <input type="checkbox"/> アルゴリズムのステップごとの動作の理解 <input type="checkbox"/> 重要な数学的な式や定理の理解 <p>週3: QAOAの実装</p> <ul style="list-style-type: none"> Day 15-17: QiskitやCirqを使用してQAOAの基本的な実装方法を学ぶ <ul style="list-style-type: none"> <input type="checkbox"/> QiskitやCirqの基本的な操作を学ぶ <input type="checkbox"/> QAOAのシミュレーションを実装を試す Day 18-21: サンプル問題を使用してQAOAの実装と最適化に取り組む <ul style="list-style-type: none"> <input type="checkbox"/> サンプル問題の選定と定義 <input type="checkbox"/> QAOAの実装と結果の評価 	<p>週4: ChatGPT (GPT-3.5) の活用</p> <ul style="list-style-type: none"> Day 22-24: ChatGPTの基本的な使用方法を学び、自然言語処理の概念を理解する <ul style="list-style-type: none"> <input type="checkbox"/> ChatGPTのインターフェイスの使い方を学ぶ <input type="checkbox"/> 自然言語処理の基本を理解 Day 25-28: ChatGPTを活用して量子コンピューティングに関する質問や課題を解決する <ul style="list-style-type: none"> <input type="checkbox"/> 質問のリストアップ <input type="checkbox"/> ChatGPTを使用して質問を解決 <p>週5: BingAI (GPT-4.0) の活用と量子最適化への応用</p> <ul style="list-style-type: none"> Day 29-31: BingAIの高度な機能と自然言語生成の学習 <ul style="list-style-type: none"> <input type="checkbox"/> BingAIの主な特長を学ぶ <input type="checkbox"/> 自然言語生成の応用例を試す Day 32-35: BingAIを使用してQAOAや量子最適化に関する複雑な問題に取り組む <ul style="list-style-type: none"> <input type="checkbox"/> 問題のリストアップ <input type="checkbox"/> BingAIを使用して問題を解決 <p>週6: プロジェクトの締めくくり</p> <ul style="list-style-type: none"> Day 36-38: 6週間の学習と実験を振り返り、進捗報告の準備 <ul style="list-style-type: none"> <input type="checkbox"/> 学習ポイントのリストアップ <input type="checkbox"/> 進捗報告のドキュメント作成 Day 39-42: インターンシッププロジェクトの成果物のまとめとプレゼンテーションの準備 <ul style="list-style-type: none"> <input type="checkbox"/> 成果物のリストアップと整理 <input type="checkbox"/> プレゼンテーションのスライド作成
--	---

図 10. ChatGPT(GPT-4.0) が作成した実習計画書

筆者はこの実習計画表に従い、インターンシップ活動を行った。

5.2 生成 AI を使った Python コードの生成

この項目については、生成 AI で Python コードを生成したりエラーコードの修正を試みたりする実験で活用した。その活用事例に関して、以降「6. 生成 AI と量子回路の生成」の中で詳しく説明していく。

6. 生成 AI と量子回路の生成

6.1 量子回路生成の実験の概要

ここでは、生成 AI を活用して、Python の外部ライブラリ Qiskit[15] を使った量子回路の生成の実験について記述する。この実験では「生成 AI を使って、量子回路の生成と数式の表示および、量子計算の質問のプロンプトに対する的確な回答を目指す」ことを目的としている。また、ChatGPT には「カスタム指示」といった機能があり、それを設定すると ChatGPT をパーソナライズ化することができる。これは ChatGPT のみがもっている機能だが、他の生成 AI サービスでも同様の指示を与えることで、カスタム指示を再現する。今回、この実験で使った生成 AI は以下である。

- ChatGPT[1]
- BingAI[3]
- Google Bard[4]

今回の実験の概要は以下のとおりである。また、今回の実験で使用した各生成 AI の種類と機能について、以下の表 1 に示す。

表 1. 今回の実験で使用した各生成 AI の種類と機能

生成AIの種類	インターネット接続	使用料	Pythonの実行	Dataのアップロード
ChatGPT(GPT-3.5) Default	×	なし	不可	不可
ChatGPT(GPT-3.5) Custom Instructions	×	なし	不可	不可
ChatGPT(GPT-4.0) With DALL-E, browsing and analysis	○	あり	可	可
ChatGPT(GPT-4.0) Plugins	○	あり	不可	不可
Bing AI(GPT-4.0)	○	なし	不可	△ 画像のみ
Google Bard(GPT-4.0)	○	なし	不可	△ 画像のみ

【今回の実験の概要と成功とみなす条件】

- Python のバージョンは 3.10.12(Google Colab ※ 2023/11/13 時点)
 - Qiskit のバージョンは 0.45.0(※ 2023/11/13 時点)
 - ChatGPT でサーバ混雑による NetworkError が多発したため、一部のプロンプトでは実験プロンプトの下に「ただし、回答は 700 文字ごとに分割し、私が「続き」と入力したら続きの回答を投稿してください。」と記述する。
 - プロンプトを送信し GPT から Python コードが返ってきたあと、次のプロンプトには原則エラー文のみを送信する。
 - 生成 AI が“ ネットなどを参照して ...etc” と返答し、Python コードを記述しなくなった場合は、コードは動かなかったものとみなし、やり取りの回数を 10 回とする。
 - 生成 AI が答えまで記述せず、中途半端な返答をしてきた場合は、その返答内にある文を使って次のプロンプトを決める。「例：「測定結果から得られた値を古典的なアルゴリズムで解析し、15 の素因数を求めます。…」のとき、「測定結果から得られた値を古典的なアルゴリズムで解析してください。』」
 - カスタム指示を使用しないのは「ChatGPT(GPT-3.5) -Default-」のみで、その他の生成 AI はすべてカスタム指示を使用する。
 - 9 回目のプロンプトを送信し、その返答にあるコードでエラーが発生した場合、そのコードは動かなかったものとみなす。
 - 生成 AI の回答が間違っている場合、次のプロンプトを「答えが違います。」とする。
- また、量子回路生成の実験の最適化のためにカスタム指示を導入した。カスタム指示は以下のとおりである。

【カスタム指示】

ChatGPT にあなたについて何を知らせれば、より良い応答を提供できると思いますか？

>> 量子計算および量子コンピューティングのスペシャリスト

ChatGPT にどのように応答してほしいですか？

>> 回答は主に **qiskit** を使う

実行環境は、**Python3.10.12** で **qiskit** のバージョンは **0.45.0** である。

全ての数式、文字式を **LaTeX** 形式で表示する。

量子アルゴリズムや量子回路について解説を行う際は、必ず **Python** コード (量子回路は **'mpl'** 表示) とアルゴリズムについて表す数式 (必ず **LATEX** 形式) を使ってわかりやすく教える。

Python コードには、1 行ごとにコメントアウトを残し、その回路やコードが何を示しているか詳しく記述する。

量子回路を実行結果で示すとともに、測定結果を保存する **"measure"** を必ず記述する。

そして、今回実験で使用するプロンプトの一覧を以下に示す。

【今回指示するプロンプト一覧】

- K_1 : 2 量子ビットの量子重ね合わせ状態について、量子回路を実装して詳しく説明してください。
- K_2 : Shor のアルゴリズムについて 15 の素因数を求める量子回路を実装し、詳しく説明してください
- K_3 : 2 量子ビットにおける量子フーリエ変換を実装し、詳しく説明してください。
- K_4 : QAOA を使って量子回路を実装し、MaxCut 問題を解いてください
- K_5 : 量子振幅推定を用いて量子数値積分を実装してください。求める解は $\sin^2(x)$ の 0 から 1 までの定積分の値です。

ここで注意点として、今回 ChatGPT(GPT-3.5) -Default- と ChatGPT(GPT-3.5) -Custom Instructions- について、比較実験のために赤字にしてある ChatGPT(GPT-3.5) -Default- では、カスタム指示を導入していない状態で実験を行う。

表 1 に示している ChatGPT(GPT-4.0) -Plugin- について、以下の 3 つを使用する。

【ChatGPT(GPT-4.0) -Plugin-】

- BrowserOp
→ウェブページの詳細情報を取得し、リアルタイムの検索エンジンの結果を取得するためのツール。
- Data Interpreter
→ SandBox 環境を使って Python コードのデータ解析を可能にするツール
- AskTheCode
→リポジトリの構造を分析し、関連するファイルの内容を取得して質問に答えるツール

さらに、今回の実験で“生成 AI のコード生成における 1 つのプロンプトに対してエラーを修正する能力”を表す指標であるエラー修正率 E_m と、“生成 AI のコード生成における実験するすべてのプロンプトのエラーを修正する能力の平均値”を表す指標である平均エラー修正率 E_{avg} の 2 つを導入した。これらの E_m と E_{avg} について、以下のルールを設け、数式とその例についても示す。

【指標のルール】

- エラーに対してどれくらい修正できるかを表す指標 (0 に近いほど修正力がある)
- 最初のプロンプトに対する回答でコードが動いた場合、 $E_m = 0$ とする
- 9 回目のやり取りの後の回答でコードが動かなければ、そのコードは動かないものとして、 $E_m = 1$ とする

【指標の数式】

“生成 AI のコード生成における 1 つのプロンプトに対してエラーを修正する能力”を表す指標であるエラー修正率 E_m を以下の数式で表す。

～ m 個目のプロンプトにおけるエラー修正率 E_m ～

$$E_m = \sin\left(\frac{\pi}{2} \cdot \frac{K_m - 1}{r_{max} - 1}\right)$$

生成 AI のコード生成における実験するすべてのプロンプトのエラーを修正する能力の平均値を以下の数式で表す。

～平均エラー修正率 E_{avg} ～

$$E_{avg} = \frac{E_1 + E_2 + \dots + E_{m=n}}{n}$$

n : 総プロンプト数, r_{max} 生成 AI とのやり取りの総回数 (ただし, $1 < r_{max}$)

K_m : m 個目のプロンプトに関して何回目のやり取りでコードが動いたか ($1 \leq K_m \leq r_{max}$)

【指標の使用例】

例) $n = 5, r_{max} = 5, E_1 = 0, E_2 = 0.38, E_3 = 1, E_4 = 1, E_5 = 0.92$ のとき

$$E_{avg} = \frac{0 + 0.38 + 1 + 1 + 0.92}{5}$$

$$E_{avg} = 0.66$$

6.2 実験結果

今回の実験において、検証を行った $K_1 \sim K_5$ と E_{avg} を以下の表 2 に示す。

表 2. 今回の実験に使用した各生成 AI の結果

生成AIの種類	インターネット接続	使用料	Pythonの実行	Dataのアップロード	→実験値	K_1	K_2	K_3	K_4	K_5	E_{avg}
ChatGPT(GPT-3.5) Default	×	なし	不可	不可		10	10	2	10	10	0.83
ChatGPT(GPT-3.5) Custom Instructions	×	なし	不可	不可		1	10	10	10	10	0.80
ChatGPT(GPT-4.0) With DALL-E, browsing and analysis	○	あり	可	可		2	2	2	10	10	0.504
ChatGPT(GPT-4.0) Plugins	○	あり	不可	不可		1	3	2	10	10	0.503
Bing AI(GPT-4.0)	○	なし	不可	△ 画像のみ		1	10	1	10	10	0.60
Google Bard(GPT-4.0)	○	なし	不可	△ 画像のみ		1	10	1	10	10	0.60

この実験結果から「ChatGPT(GPT-4.0) -Plugins-」の E_{avg} が最も小さいことから最も性能が高く、「ChatGPT(GPT-3.5) -Default-」の E_{avg} が最も大きいことから最も性能が低いといった結果になった。今回の実験のポイントとして、ChatGPT のうち GPT-4.0 を搭載している生成 AI のサービスのみが、唯一 K_2 のプロンプトの解を導くことができた点についてだ。

ゆえに総合的に判断すると、インターネット接続が可能で、画像のみならず pdf や pptx などの拡張子のファイルの解析が可能で、「ChatGPT(GPT-4.0) -Plugins-」とほぼ同水準の

性能である「ChatGPT(GPT-4.0) -With DALL-E,browsing and analysis-」が量子計算の学習において最も良いと判断する。

その一方、 K_4 と K_5 のプロンプトに対して実験に成功した生成 AI サービスはなかったため、それを改善するためにカスタム指示の内容を変えたり生成 AI に量子アルゴリズムの手法についてファインチューニングを行ったりする必要があると考える。さらに、すべての生成 AI サービスに見られた特徴として、現在の Qiskit の新しいバージョンが刷新される以前のデータのモデルであるため、現在の Qiskit のバージョンでは使えないクラスや関数が存在した。

6.3 生成 AI を活用した新しい学習法

今回の実験から、生成 AI を活用して量子計算の学習に活かすことは、今回実験で使用したプロンプトにおいて有効であるといった結果に至った。また、今回の実験で使用した各々の生成 AI の性能を、定量的な値で比較することができた。その一方、実験の結果をみると、平均エラー修正率の数値が最も良くても 0.503 であることや、使用したプロンプト数が 5 つであるため、あらゆるプロンプトに対する検証が行われていないという問題点がある。そのため現時点では、量子計算に関するすべての質問に対して、正しい答えを出力する保証がないため、量子計算を知らない初学者が生成 AI のみを活用して量子計算を学習することは難しい。そこで生成 AI のみでの学習では不十分なことから、今回筆者が量子計算を学習するにあたり「IBM Quantum で学ぶ量子コンピュータ [12]」を採用した。この書籍の対象として、Python プログラミングを通じて、量子コンピュータのアルゴリズムやアプリケーションを学びたいエンジニアに向けた内容となっているため、この対象に合致していると量子計算の学習に大いに役立つことができると考える。

ただ、今回の実験のように、生成 AI に対してあるアルゴリズムの Python プログラミングの生成とそのコードの修正には一定の成果がみられたため、量子計算の初学者が [12] の書籍に関して十分理解することができれば、生成 AI を活用して Python を活用した量子計算の学習を効率的に進められることができると考える。

7. 考察と結論

本稿では、量子コンピューティングにおける生成 AI の活用方法について説明した。今回のインターンシップにおいて一定の成果を得ることができたが、その一方で課題も見えた。量子計算に触れたことがない初学者に関して、生成 AI のみを使って学習することは現時点では難しいという結果に至った。量子計算に触れたことがない初学者が学習を進めていくためには、生成 AI だけでなく量子計算に関する論文や記事、動画、[12] の書籍を参考にすることが不可欠であると考え。今回のインターンシップでの生成 AI の活用を含めた量子計算の学習の手順、まとめについて以下に示す。

1. 生成 AI による実習計画表の作成

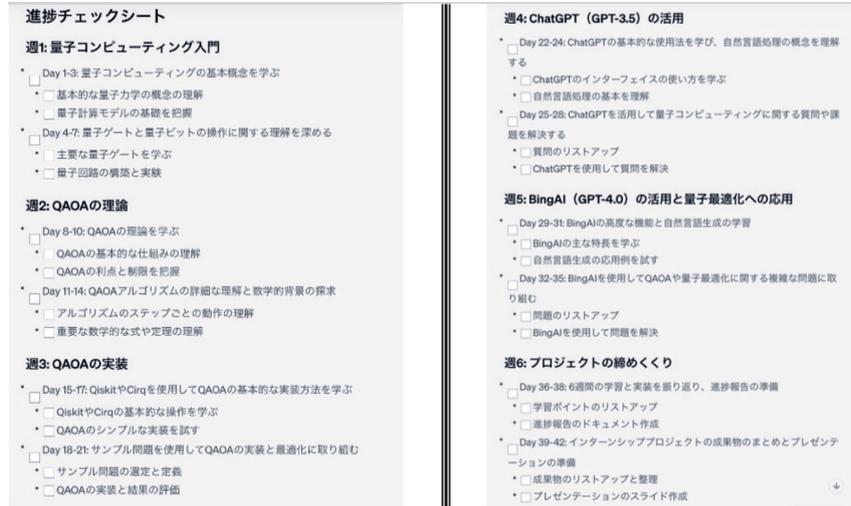


図 11. ChatGPT(GPT-4.0) が作成した実習計画書

2. [12] の書籍や記事を活用した量子コンピュータの概念および量子計算の学習
- [12] では主に、「Chapter1 量子コンピュータとは?」「Chapter2 量子コンピュータの数理」「Chapter3 IBM Quantum を使った量子計算」を学んだ。
 - インターネット上の記事では、以下の記事を使って学習した。
 - 日本総研が公開している「量子コンピュータの概説と動向」[16] 中の、第一章
 - NTT データが公開している量子コンピューティングガイドライン [17] 中の、第一章および第二章
 - IBM が公開している「Qiskit を使った量子計算の学習」[15] 中の、「1. 量子状態と量子ビット」「2. 複数量子ビットともつれ状態」
3. [12] の書籍や Youtube の動画、記事を参考にした量子アルゴリズムおよび QAOA の学習
- [12] では主に、「Chapter4 Qiskit を使った汎用量子計算」「Chapter5 量子古典ハイブリッドアルゴリズム」
 - 動画では主に以下のものを使って学習した。
 - 【誰でも量子コンピュータ！量子アルゴリズム編】 Quantum Computing for You 【第2回・9/16 実施】 [19]
 - 【誰でも量子コンピュータ！QAOA 編】 Quantum Computing for You 【追加公演第4回・9/30 実施】 [20]
 - インターネット上の記事では、以下の記事を使って学習した。
 - IBM が公開している「Qiskit を使った量子計算の学習」[15] 中の、「3. 量子プロトコルと量子アルゴリズム」「4. 量子アプリケーションのアルゴリズム」
- もちろん上記の手順が全てではないが、量子コンピュータを初めて学ぶ際は、この手順に沿って、量子計算について学習を深めていくとよい。

今回は、量子計算について生成 AI を活用した研究を立ち上げた。今回行った実験より、生成 AI が量子計算の学習において、Python プログラミングが可能で量子計算を初めて学習する人の完全なサポートを行うことができる可能性があると考え。一部の限られた問題のみではあるが、生成 AI を活用することで、量子計算の分野にも応用できることを示した。

この活用法は他の領域にも活用できると考える。その例として、ある研究を立ち上げたいときに、その内容を詳しく記述して生成 AI に研究計画表を作成してもらう。次に、生成 AI を活用してその研究に必要な文献の調査をしてもらう。ここで注意すべきこととして、生成 AI は事実に基づかない情報を生成する現象があり、これはハルシネーション [21] と呼ばれている。これを判別するために、研究を行う際は生成 AI のみに頼るのではなく、生成 AI が出力する内容について事実かそうではないかについて判断ができ、事実でない場合はそれを自力で修正できるという前提のもと、アイデアを広げる方向で活用することを推奨する。具体的にアイデアを広げた例として、今回のインターンシップで以下の図 12 のプロンプトを実行した。



図 12. “QAOA で解ける問題の例” について ChatGPT(GPT-4.0) に質問したもの

この図 12 は、筆者が QAOA で解くことのできる問題の例として、生成 AI に列挙してもらったものである。このように生成 AI の利点として、プロンプトに対して簡潔な回答かつさまざまなアイデアを提供ができるという点がある。ただ、課題としてハルシネーション [21] の現象が起こる可能性があるが、生成 AI の回答をうのみにせず、新しいアイ

アを生み出すことに活用することは、研究において活用可能であると考えられる。

この研究を通して、生成 AI の新しい活用法として、学習の計画表を作成したり文字のプロンプトからプログラミングコードを生成したり、エラーコードの修正を支援したりすることができることがわかった。しかしながら、今回の実験では十分な精度が出せていないため、さらにプロンプト数を増やしたり指示を増やしたりするなどの検証が必要である。そして、生成 AI がプロンプトに対して正しい回答と正しい構文で Python コードを出力することが、これからの課題となってくる。この生成 AI の活用は、未知数である量子計算の分野のみならず、他の未知、既知の学問でも学習の支援に役立たせることができると考えている。ゆえに、教育領域においてこのような生成 AI の活用を取り入れることで、さらなる教育環境の向上を期待している。

8. 謝辞

本稿を執筆するにあたって、インターンシップとしての活動にご協力いただいた株式会社オレンジテックラボの、西田健次 (CTO) さん、山下一郎 (技術顧問) さん、山田敏哉 (シニア UX デザイナー) さんに感謝いたします。ありがとうございました。

参考文献

- [1] OpenAI, ChatGPT, <https://openai.com/chatgpt> (accessed on 26 December 2023)
- [2] 文部科学省、初等中等教育段階における生成 AI の利用に関する暫定的なガイドライン、https://www.mext.go.jp/content/20230710-mxt_shuukyoyo2-000030823_003.pdf (accessed on 26 December 2023)
- [3] Microsoft, Bing, <https://www.bing.com/search?form=NTPCHB&q=Bing+AI&showconv=1> (accessed on 26 December 2023)
- [4] Google, Google Bard, <https://bard.google.com/chat> (accessed on 26 December 2023)
- [5] Zeeya M., "First sale for quantum computing", *nature*, Vol. 474, p. 18, (2011)
- [6] Steane, A., "Quantum computing." *Reports on Progress in Physics* 61.2 (1998): 117.
- [7] 株式会社 QunaSys, 量子コンピュータの基礎から応用まで, <https://speakerdeck.com/qunasy/quantum-summit-2019?slide=102> (accessed on 29 December 2023)
- [8] Ladd, T., Jelezko, F., Laflamme, R. et al. Quantum computers. *Nature*, Vol. 464, p.45-53 (2010).
- [9] Quantum Native Dojo, 5-3. Quantum Approximate Optimization Algorithm (QAOA): 量子近似最適化アルゴリズム, https://dojo.qulacs.org/ja/latest/notebooks/5.3_quantum_approximate_optimization_algorithm.html (accessed on 8 January 2024)
- [10] Quantum Algorithm Zoo, Navigation, <https://quantumalgorithmzoo.org/> (accessed on 8 January 2024)
- [11] Farhi, E., "Quantum supremacy through the quantum approximate optimization algorithm." *Bulletin of the American Physical Society* 62 (2017).
- [12] 湊雄一郎, 加藤拓己, 比嘉恵一朗, 永井隆太郎 『IBM Quantum で学ぶ量子コンピュータ』、秀和システム、2021年3月10日
- [13] Investor, QAOA アルゴリズムの量子回路の仕組みを理解する, <https://www.investor-daiki.com/qaoa-circuit> (accessed on 8 January 2024)

- [14] Majumdar, Ritajit, et al. "Optimizing Ansatz Design in QAOA for Max-cut." arXiv preprint at arXiv: 2106.02812, <https://arxiv.org/pdf/2106.02812.pdf>. (accessed on 8 January 2024)
- [15] Qiskit, Qiskit community, <https://qiskit.org/> (accessed on 8 January 2024)
- [16] 日本総研、量子コンピュータの概説と動向、
<https://www.jri.co.jp/MediaLibrary/file/column/opinion/pdf/11942.pdf> (accessed on 8 January 2024)
- [17] NTT データ 量子コンピューティングガイドライン、NTT DATA、
https://www.nttdata.com/jp/ja/-/media/nttdatajapan/files/news/services_info/2021/012800/012800-01.pdf (accessed on 8 January 2024)
- [18] Qiskit, 量子とは、
<https://qiskit.org/textbook/ja/what-is-quantum.html> (accessed on 8 January 2024)
- [19] 大関真之, (2022, September 16), 【誰でも量子コンピュータ！量子アルゴリズム編】 Quantum Computing for You 【第2回・9/16実施】 [video], Youtube,
https://www.youtube.com/watch?v=ZSqYDf_vSAo&t=0s (accessed on 8 January 2024)
- [20] 大関真之, (2022, September 30), 【誰でも量子コンピュータ！QAOA編】 Quantum Computing for You 【追加公演第4回・9/30実施】 [video], Youtube,
<https://www.youtube.com/watch?v=7h2mticpcFc&t=0s> (accessed on 8 January 2024)
- [21] Vipula Rawte, Amit Sheth, and Amitava Das. "A survey of hallucination in large foundation models." arXiv preprint arXiv:2309.05922 (2023).

廣峯優希 東京国際工科専門職大学 工科学部 情報工学科 3年
富山哲男 東京国際工科専門職大学 工科学部 情報工学科 教授
宮崎 淳 東京国際工科専門職大学 客員教授、株式会社オレンジテックラボ

【調査研究】

LLM の評価手法の提案と学習支援効果

安藤優希・富山哲男・山下一郎*

Proposal for LLM Evaluation Methods and Learning Support Effectiveness

Yuki Ando, Tetsuo Tomiyama and Ichiro Yamashita*

Abstract: This research study aims to evaluate both the natural language understanding and code generation capabilities of large-scale language processing quantization models, to determine their feasibility as learning aids on student PCs. The method involves using zero-shot prompting, i.e., inputting prompts without prior information, on problems posted on the competitive programming site AtCoder. The generated code will be submitted, and its accuracy and quality will be assessed. The outcome suggests that for complex problems, models tuned to instructional tasks, known as instruction models, are likely to be more effective in aiding learning.

Keywords: Large Language Models, Quantization, Artificial Intelligence

1 概要

本調査研究では、大規模言語処理量子化モデルの自然言語理解度とコード生成能力の両方を検証し、学生用 PC 上で学習の補助として使用可能なのかを調査するための研究である。手法は、競技プログラミングサイト AtCoder に掲載されている問題を zero-shot prompting、つまり、事前情報なしで prompt 入力し、生成されたコードを提出し正答率や生成されたコードから評価を定める。成果では、高度な問題であるならば命令タスクにチューニングされた instruction モデルの学習支援効果が高いと考えられる。

2 序論

今日自然言語処理の分野での進歩により、ChatGPT などの AI モデルが高度な理解能力と文章生成能力を発揮している。例えば、個人用の PC でも動作するように、メモリ使用量を減らし、推論速度を向上させても精度があまり低下しない 4bit 量子化が提案されている [8, 10, 11]。このような技術の進歩は、小中等教育や研究、ホワイトカラーの業務に新たな可能性をもたらしており、その有効活用について様々な議論が行われている [1, 2]。

* 株式会社オレンジテクラボ

また、同時に複雑なセクションを除き文書作成の支援では、人間による校正が必要ではあるが使用することができるという研究も出ている [12]。

本調査研究では、Python や java、Shell などのコードを LLM がターミナル上で ChatGPT のようなインターフェイスを介して、ローカルで実行することができる open interpreter を用い、自然言語の理解度と生成能力の両方を評価し、学習の補助に適しているモデルを調査するための研究である。また、モデルのチューニングの違いによって得意な言語や傾向が異なり、使用するライブラリによって生成されるコードが大きく変化し生成能力を正しく比較できない為、数理問題の中でも外部ライブラリを使用せずに解ける問題を取り扱い、モデルの学習データの種類によって生成されるコードの生成能力を比較し、学習の補助に最適なモデルを調査するための研究である。

3 方法論

本調査研究では、競技プログラミングサイト AtCoder の DP(動的計画法)を練習できる「Educational DP Contest / DP まとめコンテスト」から選定した問題を open interpreter を用いて、zero-shot prompting、つまり事前情報なしで prompt 入力を行いコード生成結果よりモデルの評価を行う [3]。また、推論タスクにおいてステップごとに考えるといった命令を含めることで精度が向上するが今回は使用しないことにする [9]。また、token 数が大きくなると LLM 側の指示の遵守率が低下するため、1 文字 1~3 token の日本語ではなく、基本的に 1 単語 1 token として扱う英語に翻訳しこの token 数増大による精度の低下を防ぎ、太郎君やカエルといった文章題特有の一般名詞などにモデルの理解が引っ張られないようにするため、これら一般名詞を無くした状態を prompt として扱う [4]。検証として出力されたコードを実際に AtCoder に提出することで、正答率をモデル評価に考慮する。

3.1 問題選定と Prompt 例

本調査研究では Local モデルにおける生成能力と自然言語理解の評価に重きを置いているので、評価を容易にするため DP 問題でも比較的容易な決定や計算が一連の順序付けられたステップに依存しており、DP テーブルの配列構造が単純な AtCoder の A ~ E 問題までに絞り、絶対値を用いた最小値の最適化問題である A 問題、選択肢に基づく最大値の最適化問題である C 問題、DP 問題として有名なナップサック問題である D 問題を選定した [5]。

B、E 問題では A、D 問題と同一の内容から制約のみが変更され、より難易度が高くなっているため、本調査研究での評価が不明な状態では不適切であると考えため除外した。

表 1 に本調査研究で使用した prompt 文を挙げる。

表 1：使用 prompt 文一覧

問題番号	prompt 文
A	Create Python code based on the following problem statement. Problem Statement: There are N platforms, each numbered from 1 to N, with heights denoted as h_1, h_2, \dots, h_N . Initially, you are on platform 1, and you want to reach platform N by repeating the following action multiple times: When you are on platform i, you can jump to either platform i+1 or platform i+2. When you make this jump, you incur a cost equal to the absolute difference in heights between platform i and platform j, denoted as $ h_i - h_j $.
C	Create Python code based on the following problem statement. Problem Statement: There are N days. For each $i (1 \leq i \leq N)$, you can choose one of the following options: A: Gain happiness a_i . B: Gain happiness b_i . C: Gain happiness c_i . You cannot choose the same option for two or more consecutive days. Please calculate the maximum total happiness.
D	Create Python code based on the following problem statement. Problem Statement: There are N items. Each item is numbered 1, 2, ..., N. For each $i (1 \leq i \leq N)$, item i has a weight w_i and a value v_i . You have decided to select some of these N items to take back with you. The capacity is W, and the total weight of the items you take back must not exceed W. Find the maximum total value of the items you can take back.

3.2 使用モデルについて

今回使用する local モデルは Hugging Face という機械学習モデルの開発と共有、公開をするためのプラットフォームから選定し使用する [6]。使用するモデルのフォーマットは 4bit 量子化されており、品質の低下が少ないとされている Q4_K_M の .GGUF フォーマットを用いる。また、モデルの種類は Llama2 というテキスト生成や理解 (対話ユースケース) に最適化されたモデルを、コード生成に特化させた学習データを用いチューニングした CodeLlama のベースモデル、命令タスク (自然言語の理解) に特化させたチューニングをした instruction モデル、Python に特化させたチューニングをした Python モデルの三種類の 7B、13B、34B といったパラメータ数の違うモデルを用いることとする。CodeLlama とは Meta 社が開発したコード生成専用の LLM であり、5000 億 token のプログラムに特化させた学習データを使用しているモデルのことである。また、このモデルにできることはプログラムの補完、prompt からプログラミング、プログラム関連の質問に答えるという 3 種類のことが可能である。

表 2：使用モデル例一覧

ベースモデル	instruction モデル	Python モデル
CodeLlama 7B	CodeLlama 7B	CodeLlama 7B
CodeLlama 13B	CodeLlama 13B	CodeLlama 13B
CodeLlama 34B	CodeLlama 34B	CodeLlama 34B

3.3 モデル評価手法

モデルの検証方法には、AtCoder に掲載されている「Educational DP Contest / DP まとめコンテスト」から足場という座標から移動する際のコストを最小限にする最適化問題の A 問題、選択肢に基づく最大値の最適化問題である C 問題、DP 問題として有名なナップサック問題である D 問題を用い、実際にコンテストに提出することで出た正答率とコード内容から独自の評価基準で評価する。また、提出に際して標準入力ユーザの手で加工するものとし、関数や導出処理部分に関してのみ評価の焦点をあてる。

独自の評価基準に関しては学生用 PC で動かし学習の補助ということを考慮し、「モデルの規模 (Model Size)」、「生成スピード (Generation Speed)」、「意味的正確性 (Semantic Accuracy)」、「文法的正確性 (Grammatical Accuracy)」、「問題対応力 (Problem Solving Ability)」の 5 つから見るものとする。

最大スコアを 5 とし、モデル規模はモデルごとの相対評価、生成スピードはモデルごとの絶対評価、意味的正確性や文法的正確性は減点方式の評価、問題対応力は AtCoder の正答率をもとに計算し評価する。以下に AtCoder 評価基準と独自の評価基準を用いた評価例を挙げる [7]。

3.3.1 AtCoder 評価基準

AtCoder では問題に対する解法のコードを提出することでサンプルケースが与えられ、そのサンプルケースにおいて起きた現象を以下の判定で評価する [7]。

表 3：AtCoder 判定基準一覧

AC	運営が用意したテストをパスし、正しいプログラムと判定
WA	プログラムは動いたが出力が違う
RE	コンパイル時に検出できなかったエラーが発生
CE	コンパイルエラー
IE	ジャッジシステムのエラー
TLE	指定された実行時間を超過
MLE	指定されたメモリの超過
OLE	指定されたサイズを超過

3.3.2 モデルの規模 (Model Size)

検証モデルの Model Size 種類によって値の取り方が変化する。また、この評価は 1～5 の範囲で評価される。更に、Model Size が一番小さいものを 5 としそこから相対評価をとる。本調査研究では 7B が 4.08 GB、13B が 7.87 GB、34B が 20.22 GB の三種類であり、7B を 5 とした場合、13B を 3、34B を 1 とした。

3.3.3 生成スピード (Generation Speed)

モデルの平均的な生成時間に応じて評価する。この評価もモデルの規模同様 1～5 の範囲で評価される。また、ループに陥った場合はループ発生時点までを生成スピードとして

扱うものとする。

評価基準は表4の通りである。

表4：生成スピード評価基準一覧

Time	Score
~10 秒	5
10 秒~15 秒	4
15 秒~20 秒	3
20 秒~30 秒	2
30 秒以上	1

3.3.4 意味的正確性 (Semantic Accuracy)

減点方式に関しては基礎スコアを生成コードあたりのエラー数とし、生成されたコード数を考慮し評価を変化させたいため0~1の範囲で基礎スコアに重みを付与したスコアを加算する。以下に数式例を挙げる。

$$\frac{E}{C} + \frac{E}{C} \left(1 - \frac{\log C}{\log N}\right)$$

- i. $1 - \frac{\log C}{\log N}$: コードの生成量によって変化する重み
- ii. C: 生成されたコード数 ($1 \leq C \leq N$)
- iii. E: 問題文を理解できていないコード (生成無し, 不適や全探索など) 数とする
- iv. N: 問題総数 ($1 \leq N$)
- v. コード出力なしや不適しかないもの、エラー数が極端に多く負の値になってしまう評価値は意味的正確性、文法的正確性を0評価とする。

3.3.5 文法的正確性 (Grammatical Accuracy)

意味的正確性と同様の数式を用いる。条件の変更点は以下である。

- iii. E: エラーの原因数

3.3.6 問題対応力 (Problem Solving Ability)

- i. m: WA, TLE, MLE, OLE の数
- ii. N: 問題総数とする
- iii. tは任意の数値であり、全てWAだった場合の最低スコアを表している。今回はエラー毎に評価値を変え、評価値を明瞭にするため差を均等にし、WA数の影響をなるべく受けるよう調整し0.3とする。

表 5：問題対応力評価基準

AC	1
WA/TLE/MLE/OLE	$t + (1 - t) \times (1 - \frac{m}{N})$
RE	0.2
CE	0.1
生成不適, 無し	0

3.4 AtCoder サンプル問題傾向

AtCoder にはユーザの問題への理解度や解決方法を確認するためサンプルケースが与えられ、それぞれのケースごとの出力結果によって判定が変化する。また、問題文にはプレサンプルデータが掲載されており、ユーザはこのサンプルデータを基にコードを記述していく。

3.4.1 AtCoder A 問題

準入力では足場の量 N が最初に与えられ、次に足場の高さ h が与えられる。制約は $2 \leq N \leq 10^5$ 、 $1 \leq h_i \leq 10^4$ 。

表 6：AtCoder A 問題サンプルケース傾向一覧

ケース名	サンプル問題傾向
0_00	プレサンプルデータと同一
0_01	プレサンプルデータと同一
0_02	プレサンプルデータと同一
1_00	N が小さい. h がすべて同じ値
1_01	N が小さい. h がとれる値の最少と最大
1_02	N がとれる最大の値-1. h がとれる値の最少と最大を交互
1_03	N がとれる最大の値. h がとれる値の最少と最大をパターン化
1_04	N が大きい. h はとれる値内でランダム
1_05	N が大きい. h はとれる値内でランダム
1_06	N が大きい. h はとれる値内でランダム
1_07	N が大きい. h はとれる値内でランダム

3.4.2 AtCoder C 問題

標準入力では休みの総日数 N が与えられ、 N の値に応じて各日の幸福度 3 種類 (a, b, c) が与えられる。

制約は $1 \leq N \leq 10^5$ 、 $1 \leq a_i, b_i, c_i \leq 10^4$ 。

表7: AtCoder C 問題サンプルケース傾向一覧

ケース名	サンプル問題傾向
0_00	プレサンプルデータと同一
0_01	プレサンプルデータと同一
0_02	プレサンプルデータと同一
1_00	N がとれる最少の値. a, b, c がすべてとれる最大の値
1_01	N とすべての a, b, c が, とれる最大の値
1_02	N が大きい. a, b, c はとれる値内でランダム
1_03	N が大きい. a, b, c はとれる値内でランダム
1_04	N が大きい. a, b, c はとれる値内でランダム
1_05	N が大きい. a, b, c はとれる値内でランダム
1_06	N が大きい. a, b, c はとれる値内でランダム
1_07	N が大きい. a, b, c はとれる値内でランダム
1_08	N が大きい. a, b, c はとれる値内でランダム
1_09	N が大きい. a, b, c はとれる値内でランダム

3.4.3 AtCoder D 問題

標準入力ではものの数 N と重さの総和 W が与えられ、 N の値に応じて順次重さ w と価値 v が与えられる。

制約は $1 \leq N \leq 100$ 、 $1 \leq W \leq 10^5$ 、 $1 \leq w_i \leq W$ 、 $1 \leq v_i \leq 10^9$ 。

表8: AtCoder D 問題サンプルケース傾向一覧

ケース名	サンプル問題傾向
0_00	プレサンプルデータと同一
0_01	プレサンプルデータと同一
0_02	プレサンプルデータと同一
1_00	N がとれる最少と W がとれる最大の値. w, v がとれる最大の値
1_01	N, W, w, v がとれる最大の値
1_02	N がとれる最大の値. W, w, v はとれる値内の大きい値でランダム
1_03	N がとれる最大の値. W, v はとれる値内の大きい値でランダム. w はとれる値内でランダム
1_04	N がとれる最大の値. W, v はとれる値内の大きい値でランダム. w はとれる値内でランダム
1_05	N がとれる最大の値-1. W, v はとれる値内の大きい値でランダム. w はとれる値内でランダム
1_06	N がとれる最大の値. W, v はとれる値内の大きい値でランダム. w はとれる値内でランダム
1_07	N がとれる最大の値-1. W, v はとれる値内の大きい値でランダム. w はとれる値内の小さい値でランダム
1_08	N がとれる最大の値-1. W, v はとれる値内の大きい値でランダム. w はとれる値内の小さい値でランダム
1_09	N がとれる最大の値. W, v はとれる値内の大きい値でランダム. w はとれる値内の小さい値でランダム

4 結果

4.1 出力されたコード例

4.1.1 AtCoder A 問題

全体的に絶対値の扱い方をよく理解できていないのか、図2のように if 分岐の下で行われている絶対値の加算が同一の計算をおこなっているなど、条件分岐が意味をなしていないものや図1のように計算に必要な情報を利用していないものが見受けられた。

```
def jumping_on_platforms(heights):
    # Initialize variables
    n = len(heights)
    dp = [[0] * (n + 1) for _ in range(2)]

    # Fill in the table
    for i in range(1, n + 1):
        for j in range(i + 1, n + 1):
            dp[i % 2][j] = min(dp[(i - 1) % 2][j], dp[(i - 2) % 2][j - 1])

    # Return the optimal cost
    return dp[n % 2][n]
```

図1：CodeLlama 34B instruction

```
import sys

def jumping_cost(heights):
    cost = 0
    for i in range(len(heights) - 1):
        if heights[i] > heights[i + 1]:
            cost += abs(heights[i] - heights[i + 1])
        elif heights[i] < heights[i + 1]:
            cost += abs(heights[i + 1] - heights[i])
    return cost
```

図2：CodeLlama 13B instruction

4.1.2 AtCoder C 問題

ほとんどのモデルでコードの出力が不十分であった。また、図3のように DP テーブルを作成しているにも関わらず、その後の条件分岐で制約を無視したコードや、図4のように最初から制約を無視するなど、制約の無視が多く見受けられた。

```

1 def max_happiness(a, b, c):
2     n = len(a)
3     dp = [[0] * (n + 1)] * (n + 1)
4     prev = [-1] * (n + 1)
5     dp[0][0] = a[0]
6     prev[0] = 0
7
8     for i in range(1, n):
9         option_a = dp[i - 1][i - 1] + a[i]
10        option_b = dp[i - 2][i - 2] + b[i]
11        option_c = dp[i - 3][i - 3] + c[i]
12
13        if option_a >= option_b and option_a >= option_c:
14            dp[i][i] = option_a
15            prev[i] = 0
16        elif option_b >= option_a and option_b >= option_c:
17            dp[i][i] = option_b
18            prev[i] = 1
19        else:
20            dp[i][i] = option_c
21            prev[i] = 2
22
23    return dp[n - 1][n - 1]
24

```

図 3 : CodeLlama 34B instruction

```

# Define the happiness values for each option
happiness_A = [a_1, a_2, a_3, a_4, a_5]
happiness_B = [b_1, b_2, b_3, b_4, b_5]
happiness_C = [c_1, c_2, c_3, c_4, c_5]

# Define the maximum total happiness
max_total_happiness = 0

# Iterate through each day and calculate the maximum total happiness
for i in range(N):
    # Calculate the maximum total happiness for each day
    max_total_happiness = max(max_total_happiness, happiness_A[i] + happiness_B[i] + happiness_C[i])

# Print the maximum total happiness
print("The maximum total happiness is", max_total_happiness)

```

図 4 : CodeLlama 7B instruction

4.1.3 AtCoder D 問題

DP 有名問題ということもあってか、DP 法を使用したコードが多く存在した(図 5)。また、DP 法を使っていないコードは図 6 のように、制限まで取得し続けるものや重さと価値の比率で取得するものであった。

```
def knapsack(weights, values, capacity):
    # Initialize variables
    n = len(weights)
    dp = [[0] * (capacity + 1) for _ in range(n + 1)]

    # Fill in the table
    for i in range(1, n + 1):
        for j in range(1, capacity + 1):
            if weights[i - 1] > j:
                dp[i][j] = dp[i - 1][j]
            else:
                dp[i][j] = max(dp[i - 1][j], dp[i - 1][j - weights[i - 1]] + values[i - 1])

    # Return the optimal value
    return dp[n][capacity]
```

図 5 : CodeLlama 34B instruction(DP)

```
import sys

def get_max_value(weights, values, capacity):
    # Initialize variables
    max_value = 0
    current_weight = 0
    current_value = 0

    # Sort the items by value/weight ratio in descending order
    sorted_items = sorted(zip(values, weights), key=lambda x: x[1]/x[0], reverse=True)

    # Iterate through the items and add them to the knapsack until the capacity is exceeded
    for value, weight in sorted_items:
        if current_weight + weight <= capacity:
            max_value += value
            current_weight += weight

    return max_value
```

図 6 : CodeLlama 13B instruction(比率)

4.2 それぞれのモデル差異

4.2.1 ベースモデル

コード出力は2つのモデルしか出力せず、4.2.1.1 CodeLlama 13B(A 問題)のように問題の意図を理解していないものと4.2.1.2 CodeLlama 13B(D 問題)のように理解しているものと極端な生成が見受けられ不安定であった。

```
def f(n):
    return n * (n + 1) / 2
```

図 7 : CodeLlama 13B(A 問題)

```

def max_value(W, wt, val):
    n = len(wt)
    K = [[0 for x in range(W + 1)] for x in range(n + 1)]
    for i in range(n + 1):
        for w in range(W + 1):
            if i == 0 or w == 0:
                K[i][w] = 0
            elif wt[i - 1] <= w:
                K[i][w] = max(val[i - 1] + K[i - 1][w - wt[i - 1]], K[i - 1][w])
            else:
                K[i][w] = K[i - 1][w]
    return K[n][W]

```

図 8 : CodeLlama 13B(D 問題)

4.2.2 instruction モデル

他モデルよりも DP 法らしき手法をとっているものが多かったが、図 3 のように DP テーブルの初期化不足や不必要な変数の使用、図 9 のように if 分岐後の処理に範囲外のインデックスを選択する可能性を残していたりなど初歩的なミスが多く見受けられた。

```

def min_cost(heights):
    if len(heights) == 0:
        return 0
    elif len(heights) == 1:
        return 0
    else:
        cost1 = abs(heights[0] - heights[1]) + min_cost(heights[2:])
        cost2 = abs(heights[0] - heights[2]) + min_cost(heights[1:])
        return min(cost1, cost2)

print(min_cost([h_1, h_2, ..., h_N]))

```

図 9 : CodeLlama 7B instruction(A 問題)

4.2.3 Python モデル

Python モデルでは具体的なコードは出力せず、4.2.3.1 CodeLlama 7B Python(D 問題) のように関数の型だけ書いて終わるものがほとんどであった。

```

def max_value(W, w, v):
    # TODO: write function here
    return None

```

図 10 : CodeLlama 7B Python(D 問題)

4.3 AtCoder による評価結果

表 9：AtCoder による評価結果一覧

問題/モデル名	CodeLlama 7B	CodeLlama 13B	CodeLlama 34B	CodeLlama 7B inst	CodeLlama 13B inst	CodeLlama 34B inst	CodeLlama 7B Python	CodeLlama 13B Python	CodeLlama 34B Python
A	RE			RE	7	9			
C				13	RE	10			
D		0		RE	8	0			
AC									
WA									
RE									
TLE									
出力無し									

※表内の数値は AC 以外の数

以上の結果から C 問題がもっとも正答率、生成率が低いことがわかり、A 問題は多様な判定結果が見受けられる。また、D 問題は生成されたコードの正答率は他問題よりも高い傾向にあった。

4.4 独自による評価結果

表 10：独自による評価結果一覧

	CodeLlama 7B	CodeLlama 13B	CodeLlama 34B	CodeLlama 7B inst	CodeLlama 13B inst	CodeLlama 34B inst	CodeLlama 7B Python	CodeLlama 13B Python	CodeLlama 34B Python
Model Size	5.0	3.0	1.0	5.0	3.0	1.0	5.0	3.0	1.0
Generation Speed	5.0	4.0	4.0	5.0	3.0	3.0	5.0	5.0	5.0
Semantic Accuracy	1.0	3.631	0.0	3.666	3.631	4.667	0.0	0.0	0.0
Grammatical Accuracy	3.0	5.0	0.0	4.333	4.315	3.667	0.0	0.0	0.0
Problem Solving Ability	0.2	1.0	0.0	0.7	1.124	1.889	0.0	0.0	0.0

以上の結果から汎用的なベースモデルには意味的正確性に大きなばらつきがみられた。また、文法的正確性に関してはどちらもばらつきがみられ、そこまでコード生成能力に差が見られないと考えられる。問題対応力では平均して instruction モデルが高い水準を誇っていた。

図 11 にレーダーチャートを表示する。

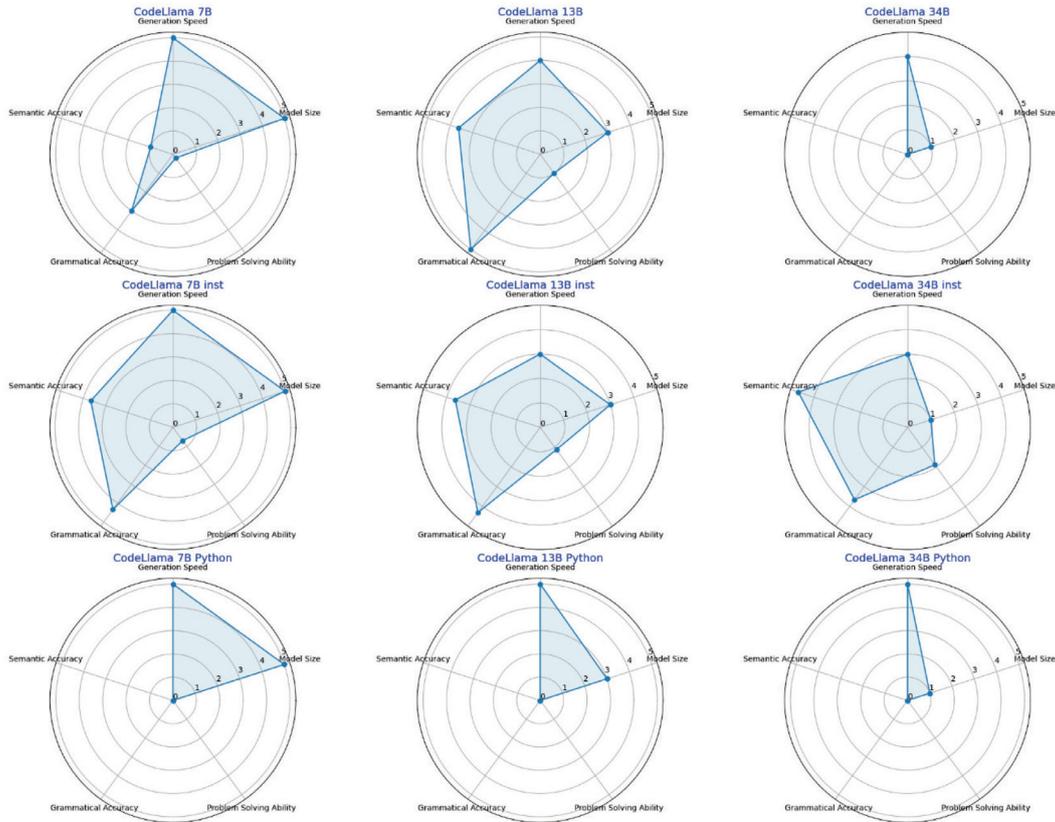


図 11：独自評価による結果をレーダーチャートにしたもの

5 考察

AtCoder による評価結果では、instruction モデルのような自然言語理解に特化させたモデルにおいて、ナップサック問題といった複雑で有名な問題は学習データに含まれている可能性があり、平均的な生成精度が高いのだと考えられる。

また、instruction モデルでは平均的な生成量は高いが正答数を見てみるとベースモデルに比べ、生成精度がモデルサイズに依存していることがわかる。このことから instruction モデルでは他モデルよりもパラメータ数が生成精度に大きな影響を与えていると考えられる。

独自の評価結果では、AtCoder と違った評価を得ることができた。例えば同規模の CodeLlama 13B と CodeLlama 13B instruction では、表 9 を見比べるとベースモデルの CodeLlama 13B の方が完全正答を出しており、instruction モデルと比べ生成量も同じであり間違えた数も少ないが、ベースモデルは完全正答と全不正解と二極化した結果となった。しかし、instruction モデルでは正答率こそ低いものの問題別に同じような正答率を出していることからモデルの安定性や汎化能力という面では instruction モデルが優れていると考える。そのため、表 5 はこの基準でも適切な評価を得られたと考える。

また、意味的正確性や文法的正確性はユーザの問題やコードへの理解度やユーザの判断

基準に依存してしまうため、判定基準を検証ごとに定める必要があると考える。更に、エラーの重大度などを考慮していないためこの評価基準は見直しが必要であると考えられる。

最後に、今回の研究では Python モデルの出力はすべての問題を通して図 10 のように出力不適、または無しであった。しかし、機械学習のテンプレート作成などにおいては import してくるモジュールなどに違いがあったため、数理アルゴリズムというよりも機械学習コードの作成や添削に特化させているモデルであると考えられる。

6 結論

モデル差が見受けられたものの、一部のモデルでは学習の補助としては使えると考える。また、自然言語理解に特化させた instruction モデルではほとんどのモデルが何かしらのコードを出力しており、平均的なモデル精度は比較的高かった。このことから、数理解習の補助に関しては instruction モデル且つパラメータ数の多いモデルのほうが良いと考える。

また、モデルの精度が非常に高いものは見受けられなかったが、ある程度のテンプレートは作成されていたので、被学習者が正しい知識を保持していれば生成コードを参考に添削可能であった。

7 謝辞

本研究を進めるにあたり、貴重な指導と実験に関するアドバイスを下さった西田健次 CTO 様、山田シニア UX デザイナー様、宮崎 淳 東京国際工科専門職大学 客員教授、オレンジテックラボ CEO 様に深く感謝いたします。

参考文献

- [1] 生成 AI の利用について：文部科学省 (mext.go.jp), 2023 年 7 月 4 日 . accessed December 10, 2023.
https://www.mext.go.jp/a_menu/other/mext_02412.html
- [2] 経済産業省 . 生成 AI 時代の DX 推進に必要な人材・スキルの考え方 . デジタル時代の人材政策に関する検討会 , 2023 年 8 月 7 日 . accessed December 10, 2023.
<https://www.meti.go.jp/press/2023/08/20230807001/20230807001.html>
- [3] Educational DP Contest / DP まとめコンテスト , accessed October 23, 2023.
<https://atcoder.jp/contests/dp>
- [4] Anthropic Claude2 で Prompt 実行の精度を高めるテクニック , accessed October 30, 2023.
<https://qiita.com/nkitaarashi/items/85ac0eb7c451f30972d4>
- [5] Educational DP Contest の F ~ J 問題の解説と類題集 , accessed December 26, 2023.
<https://qiita.com/drken/items/03c7db44ccd27820ea0d>
- [6] hugging face の紹介 , accessed December 28, 2023.
<https://zenn.dev/yoshikawat64m/articles/6c7862ad376368>
- [7] 競技プログラミング用語集 , accessed December 28, 2023.
<https://atcoder.jp/contests/abc074/glossary?lang=ja>

- [8] LIU, Zechun, et al. LLM-QAT: Data-Free Quantization Aware Training for Large Language Models. In International Conference on Learning Representations (ICLR), September 19, 2023. accessed January 18, 2024.
- [9] KOJIMA, Takeshi, et al. Large language models are zero-shot reasoners. Advances in neural information processing systems, 2022, 35: 22199-22213. accessed January 18, 2024.
- [10] LIU, Jing, et al. QLLM: Accurate and Efficient Low-Bitwidth Quantization for Large Language Models. In International Conference on Learning Representations (ICLR), January 16, 2024. accessed January 18, 2024.
- [11] FRANTAR, Elias, et al. OPTQ: Accurate Quantization for Generative Pre-trained Transformers. In International Conference on Learning Representations (ICLR) poster. February 02, 2023. accessed January 19, 2024.
- [12] MICHELET, Gaëtan; BREITINGER, Frank. ChatGPT, Llama, can you write my report? An experiment on assisted digital forensics reports written using (Local) Large Language Models. In Digital Forensics Research Conference (DFRWS EU). December 29, 2023. accessed January 19, 2024.

安藤優希 東京国際工科専門職大学 工科学部 情報工学科 3年
富山哲男 東京国際工科専門職大学 工科学部 情報工学科 教授
山下一郎 株式会社オレンジテクラボ

『東京国際工科専門職大学紀要』規程

(名称及び発行)

第1条 東京国際工科専門職大学紀要と称し、英語名を Bulletin of International Professional University of Technology in Tokyo とする。

(委員会の設置)

第2条 東京国際工科専門職大学は紀要を発行するため紀要委員会（以下「委員会」という）を設置する。

2 委員会は編集委員長および編集委員若干名から構成される。

(投稿者・分野)

第3条 投稿者は東京国際工科専門職大学の常勤あるいは非常勤の教員、助手、ティーチングアシスタント、その他学部長が認めた者とし、分野は各研究分野を主たる対象とする。共著者は他の研究機関等に所属する者も可能とする。

2 投稿された原稿の分類は、論文、総説、研究ノート、調査報告、研究資料等とする。

(執筆要項)

第4条 投稿は、別に定める執筆要項に基づかなければならない。

(査読及び採録)

第5条 投稿された原稿は、論文については原則的に査読を受ける。この結果を踏まえて委員会は、全ての原稿の採録可否の決定及び編集を行う。

2 査読の方法は別に定める。

3 査読者は1名以上とする。

(投稿原稿について)

第6条 投稿する論文原稿は、未発表のものに限り、他の学術誌等にすでに掲載されたもの、または投稿中のものと同一あるいは類似した内容の原稿は掲載できない。ただし、総説、研究ノート、調査報告、研究資料等はこの限りではない。また、学会等での口頭発表の原稿に加筆修正を加えたものはその旨を記載しなければならない。

2 一度提出した原稿は、発行までの過程において、大幅に変更することはできない。

(発行及び公開)

第7条 発行は原則として年に1回とする。

2 公開は紙媒体や電子媒体とする。

(著作権等)

第8条 原稿の著作権は、投稿者（著者）に帰属する。他誌などにその全部または一部を使用する場合には、出典が本紀要であることを明記する。

- 2 前項のうち複製権、公衆送信権については、東京国際工科専門職大学に許諾する。
- 3 第三者の有する知的所有権については、投稿者の責任において、著者・発行者等の知的所有権者と交渉して当該問題を処理し、各種権利処理の状況を編集委員会に報告する。

この規程は、2023年5月15日より施行する。

『東京国際工科専門職大学紀要』執筆要項

原稿の長さ

1. 原稿の長さは原則として以下の範囲内とする。なお、文字数には要約、注、参考文献、図表等を含める。制限字数を大幅に超えたものは受理しない場合がある。
 - ・ 論文、総説：20 ページ以内
 - ・ 研究ノート：20 ページ以内
 - ・ 調査報告：15 ページ以内
 - ・ 研究資料：15 ページ以内

要約とキーワード

1. 論文には英文要約（原則 300 語以内）とキーワード（原則日本語または英語 3～5 語程度）を置く。

書式

2. 執筆原稿は日本語または英語とする。
3. ファイルを提出する形式は原則として pdf とする。
4. 原稿は A4 版、1 頁全角文字の場合 40 字×40 行とした 1,600 字が最大となる、横書きを原則とする。
5. 論文は表題（日本語と英語）、英文要約、キーワード、本文、注、参考文献の順とする。論文以外の原稿は表題、本文、注、参考文献の順とする。
- 6.. 図表は原稿の該当箇所にあらかじめ挿入しておく。図は下にキャプションを付ける。表は上にキャプションをつける。
- 7.. フォントおよびフォントサイズは下記の通りとする。
- 8.. 句読点は、和文は、「、」「。」、英文は、「,」「.」とする。

	表題	著者名	要約と キーワード	本文	注と 参考文献
和文	和文： ゴシック 18pt 英文： Times 14pt	和文： 明朝 14pt 英文： Times 12pt	和文： 明朝 10.5pt 英文： Times 10.5pt	明朝 10pt	明朝 9pt
英文	ゴシック 18pt	Times 14pt	Times 12pt	Times 10pt	Times 10pt

投稿方法

1. 投稿の締め切りは原則として毎年 11 月末とする。
- 第 1 条 依頼原稿を除き、投稿者は東京国際工科専門職大学の常勤あるいは非常勤の教員、助手、ティーチングアシスタント、指導教員のもとの学生、その他学部長が認めた者とし、分野は各研究・教育分野を主たる対象とする。著者は東京国際工科専門職大学に属さない者も可能とする。
2. 投稿者は、紀要委員会が配布する投稿申込書に必要事項を記入し、原稿と併せて提出

する。

3. 投稿申込書および原稿の投稿は原則として PDF での提出とする。
4. 投稿された原稿は、論文については査読を受ける。この結果を踏まえて委員会は、全ての原稿の採録可否の決定及び編集を行う。
5. 投稿する原稿は、未発表のものに限り、他の学術誌等にすでに掲載されたもの、または投稿中のものと同じあるいは類似した内容の原稿は掲載できない。また、学会等での発表の原稿に加筆修正を加えたものはその旨を記載しなければならない。
6. 一度提出した原稿は、発行までの過程において、大幅に変更することはできない。

校正

1. 校正は原則として 2 回までとし、著者が責任をもって行うものとする。
2. 査読の結果、論文の修正・再提出がある場合、執筆者は指定の期日までに修正箇所を明記する形で論文を再提出する。
3. 原稿が発行過程で著しく遅延した場合、次号への投稿扱いとする。

著作権

1. 第三者の有する知的所有権については、投稿者の責任において、著者・発行者等の知的所有権者と交渉して当該問題を処理し、各種権利処理の状況を編集委員会に報告する。
2. 複製権、公衆送信権を東京国際工科専門職大学に許諾するものとする。

執筆者一覧

富山 哲男	工科学部	
武本 充治	工科学部	
三宅 茂樹	工科学部 情報工学科	AI 戦略コース
爰川 知宏	工科学部 情報工学科	IoT システムコース
兒玉 賢史	工科学部 情報工学科	IoT システムコース
辻野 雅之	工科学部 情報工学科	IoT システムコース
藤井 竜也	工科学部 情報工学科	IoT システムコース
山本 裕	工科学部 情報工学科	IoT システムコース
水上 憲明	工科学部 情報工学科	ロボット開発コース
菅谷 孝義	工科学部 情報工学科	
小野 憲史	工科学部 デジタルエンタテインメント学科	ゲームプロデュースコース
水上 恵太	工科学部 デジタルエンタテインメント学科	ゲームプロデュースコース
山口 直彦	工科学部 情報工学科	
宮崎 淳	株式会社オレンジテクラボ	
山下 一郎	株式会社オレンジテクラボ	
伊勢みゆき	学校法人江副学園新宿日本語学校	
青山 龍	ソニー株式会社	
安彦 剛志	ソニー株式会社	
八木 泉	ソニー株式会社	

学生 安藤優希、内埜大斗、廣峯優希

編集委員

辻野 雅之
小野 憲史

東京国際工科専門職大学 紀要 第3号

2024年3月31日 発行

発行 東京国際工科専門職大学
東京都新宿区西新宿 1-7-3
電話 03-3344-5555 (代表)

印刷 日本アスペクトコア株式会社
東京都新宿区新小川町 6-29