Modelling of decision-making process by inverse Bayesian inference

逆ベイズ推定を用いた意思決定プロセスのモデル化

July 2018

Youichi HORII

堀井　洋一

Modelling of decision-making process by inverse Bayesian inference

逆ベイズ推定を用いた意思決定プロセスのモデル化

July 2018

Waseda University
Graduate School of Fundamental Science and Engineering
Department of Intermedia Studies,
Research on Biological Expression

Youichi HORII

堀井　洋一

Preface and acknowledgement

Suppose you are betting on a coin toss. Which do you bet on next after the front has continued for a while? Since it is actually an independent event, the probability is 1/2. However, humans frequently bet on the front thinking "Things come in threes twice" or bet on the back thinking "Third time lucky". In order to model various decision processes of human beings and organisms, many studies using Bayesian inference in particular have been conducted. In Bayesian inference, a posteriori probability $P(h|d)$ is obtained by using the prior probability $P(h)$ of a model h having a likelihood $P(d|h)$ where a certain value d is selected and then the prior probability is updated. In other words, the Bayesian inference models the idea that "Things come in threes". However, humans also consider "Third time luck" and often make the opposite judgment from the past experience. Humans continue to make decisions in this uncertain situation where conflicting grounds are mixed. Therefore, modeling by Bayesian inference is fundamentally insufficient. On the other hand, Professor Yukio-Pegio Gunji proposed a *inverse Bayesian inference* that updates the likelihood $P(d|h)$ of model h that was fixed in Bayesian inference for each event [1]. This effectiveness has been confirmed by modeling behaviors of living creatures. This doctoral thesis summarizes the methodology of modeling human decision-making process using *inverse Bayesian inference* and the results demonstrated by experiment and various data.

Professor Gunji arrived at Faculty of Science, Kobe University in 1987, when I was studying the solar system formation process [2] [3] [4]. I received a strong impact on his extraordinary character and way of thinking. After that, I studied under Professor Gunji and compiled a master's thesis on cognitive science [5]. Since 1990, I have studied computer music [6], computer graphics [7], and human interaction [8] at corporate laboratories. Since 2010, I have been working on research topics on infrastructure and social science. I was seeking methods to deal with events involving many decision makers. Especially in infrastructure construction, uncertainty is extremely high, and various stakeholders make decisions with their respective motives. In order to grasp the situation objectively, I tried to construct a methodology incorporating visualization using computer graphics technology and interaction technology to aid understanding of data. However, humans frequently perform unreasonable judgment. Under such circumstances, how should we form consensus? Very fortunately for me, Professor Gunji was appointed to Waseda University (2.5 km from my office!) in 2014, and since September 2015 I began studying together for the first time in 25 years. In the doctoral course of Waseda University, I received huge guidance on not only the *inverse Bayesian inference* of this doctor thesis but also how to capture various kinds of *human*, *environment* and *computation*.

Contents

Chapter 1 Introduction

1.1 Modelling of decision making process by Bayesian inference

Human beings perceive and behave based on uncertain information and sometimes make unreasonable judgments when viewed in objective terms. Modeling of the process used for instantly making some kind of decision—even if unreasonable—from uncertain information for which conditions can suddenly change is important not only for understanding the human cognitive process but also for constructing intelligent systems that can adapt to actual environments. Mach [9] and Helmholtz [10] [11] suggested that *stochastic inference* provides a basis for the human cognitive process based on uncertain information. Additionally, advances in artificial intelligence research have driven research of cognition and behavior based on such stochastic inference [12] [13] [14] [15]. Given that a causal relation between two events, such as "if p then q" is true, the contraposition "if not q then not p" is true, but the converse "if q then p" or the converse of the contrapositive "if not p then not q" is not necessarily true. It is known, however, that humans will sometimes perceive the converse to be true and make unreasonable judgments as a result. This lies in the tendency to make any asymmetry in cause-and-effect relations symmetric and is therefore called "symmetry bias," which has come to be studied in detail through standardized experiments [16] [17] [18] [19]. In these experiments, subjects were asked to evaluate the extent to which the casual relation of presented events could be relied upon and attempts were made to model the processes involved. For example, the contingency model of Jenkins [20] and the probabilistic contrast model of Cheng [21] have been proposed. In addition, Hattori discovered that the causal inference parameter $(P(p|q)P(q|p))^{1/2}$ that deals symmetrically with "if p then q" and "if q then p" had a high correlation with this sense of reliability [19].

Gigerenzer showed by experiment that human thinking is performed along the lines of *Bayesian inference*, and that it can be influenced by converting expressions for rate of event occurrence into "frequency" or "probability" [22]. Meanwhile, Knill described how humans and other living organisms perform stochastic inference based on uncertain information obtained from the real world and reported that many experimentally observed examples could be explained by a Bayesian perceptual system [23]. In contrast to reacting suddenly to particular input, this system integrates information that propagates over space and time. He explained that this property holds not only for inputted information but also for uncertainty in the results of behavior. In addition, Manktelow [24] classified and analyzed how humans perceive a variety of uncertain events such as coin-toss games, probability that a weather report is accurate, existence of the Loch Ness

Monster, etc. It was described here that Bayesian inference, which is used to update a person's hypothesis after observing an event as posterior probability, is the basis for human intuitive cognition. There are also many studies that children's developmental process is in line with Bayesian principles [25], [26], [27], [28]. Pellicano, meanwhile, described how the perceptual process of autistic patients follows Bayesian inference and that autistic people tend to dislike unconventional stimuli that do not agree with experience [29]. This result suggests the ability and limitations of the Bayesian inference.

As mentioned above, the Bayesian inference has become widely used to model various cognitive and decision-making processes and hundreds of studies have been done [30] [31] [32] [33] [34]. However, Knill pointed out that Bayesian inference does not respond sensitively to unconventional input [23]. As described later in the **Chapter 3**, simple Bayesian inference does not immediately follow a sudden change in data such that the model itself changes. Meanwhile, Gallistel et al. have proposed a method "Bayesian change point analysis" which detects sudden change of data [35], [36], [37], [38]. First, for the null hypothesis that data does not change abruptly and alternative hypotheses, calculate the ratio of likelihood, that is, the Bayes factor. Using the sign of the logarithm of the Bayes factor (i.e., weight of the evidence), it is determined whether or not there is a change in the data. When it is evaluated that there is a change, the time at which the distance between the line connecting the start point and the end point of accumulated data and the cumulative value becomes maximum is set as the change point. This method has been applied to the analysis of learning curves of infants and animals. But their Bayesian approach is not an attempt to model the cognitive process itself. However, psychological experiments sometimes result in the opposite result of Bayesian inference [39], and modeling of such *Anti-Bayesian* is also attempted [40].

1.2 Inverse Bayesian inference

In response to the above, Arecchi has proposed *inverse Bayesian inference* which refers to inverse transposition of terms in Bayes formula [41]. Independent of Arecchi, Gunji et al. also proposed *inverse Bayesian inference* which refers to replacing conditional probability by marginal probability [1] [42] [43]. In this thesis, I refer to *inverse Bayesian inference* in Gunji's sense. While Bayesian inference predicts beforehand the probability that a hypothesis will be selected to directly derive an optimal solution, *inverse Bayesian inference* determines the likelihood of a hypothesis afterwards to alleviate previously prescribed rules.

The problem of perception and decision making can be thought of in the framework of *endophysics* such as understanding the nature of the world by the internal observer [44]. *Retrocausality* is a way of thinking about opposite causality and time [45]. Bayesian inference corresponds to the forward direction of time, while the *inverse Bayesian inference* corresponds to the reverse flow of time. When considering human reasoning ability and decision-making, the logic which is normally thought not to depend on time has deep connection with time [46] [47] [48]. Sawa proposed a *double Homunculus model* in which two differential equations influenced each other and expressed the recursive structure of the observer. By combining this model, various kinds of logic gates can be constructed, and the reverse logic gate allows us to consider of the reverse flow and uncertainty of time [49] [50]. In order to realize this, the *double Homunculus model* adopts hidden and dummy variables. In BIB inference, these variables correspond to likelihood and prior probability.

1.3 Effects of surrounding opinions

On the other hand, humans are often affected by surrounding opinions. When making a decision, a bandwagon effect synchronizing with the surroundings and an underdog effect which rebounds to the surroundings are known. These effects are confirmed by the influence of opinion polls before the election [51], financial transactions such as stocks and currency exchange [52], decision making at the time of introducing new technology of the company [53], and so on. Also, psychological experiments have verified the influence of majority factions on minorities [54].

Regarding the bandwagon effect and the underdog effect in elections, Simon formulated the relationship between opinion poll prediction and actual voting behavior [55]. In addition, many virtual voting experiments to verify these effects have been done, and the bandwagon effect has been confirmed [56] [57] [58] [59] [60]. *Pivotal voter model* based on the hypothesis that voters make voluntary and rational choices has been proposed. In this model, the expected value is calculated from the cost for voting, the degree of influence given by the vote, and the profit from the voting result, and whether or not to vote is judged by this expected value [61]. On the other hand, the *strategic elites model* focuses on behaviors such as candidates, political parties, and leaders of profit organizations [62]. In recent years, with the widespread use of SNS, the information obtained by voters is increasing. The results of public opinion polls are said to affect the voting behavior, and some countries restrict public disclosure during the election period [51]. Also, not only in the US presidential elections in 2000 and 2016, but also in other countries, the case of *close battle* in important elections is increasing.

Nurkse [63] had earlier warned about the dangers of the bandwagon effect that would cause instability in financial markets [52]. Allen et al. analyzed the foreign exchange trends forecasted by the *chartists* who make decisions based on past experience, not economic theory [64]. When the bandwagon effect is seen, the elasticity of expectations for exchange rate fluctuation increases and the market becomes unstable. Siddiqi showed that the broker can control equilibrium state of stock price (perfect Bayesian equilibrium) in Bayesian game by three people in institutional investors, individual investors and brokers in the stock market [65] [66].

Even when companies adopt new technology, it is known that the bandwagon effect can be seen [53]. Decision making of technology introduction is influenced not only by reasonable judgment such as technical or profitability (i.e. efficient-choice theories) but also by the number and reputation that has been adopted in advance (i.e. fad theories) [67] [68].

Many psychological studies have examined the tendency of humans to synchronize with surrounding situations. Asch conducted an experiment to answer the length of the line segment drawn on the card in a group with seven confederates for each subject [54]. All seven confederates return the same wrong answer, and finally the subject answers. It has been shown that at least 75% tend to synchronize with confederates' answers at least once and tend to be in line with the majority opinion. Moscovici, on the other hand, examined the influence of the minority on the majority [69]. In the experiment, four subjects and two confederates answer the color of the projected slide. Even if the minority's confederates are wrong, when the answers are consistent, it has become clear that they affect the majority. Moreover, the influence tends to be strong when the way of thinking is flexible [70] [71] and the social attributes are similar [72].

There are a few studies that associate the bandwagon effect with Bayesian inference. Rosenkopf et al. pointed out that the bandwagon effect was observed when adopting new technology, the criteria were updated according to the circumstances, and that the process was a Bayesian mechanism [53]. Chappell et al. proposed a *Bayesian Bandwagon Model* [73]. They analyzed data on the decision process of the federal funds rate by members of the Federal Open Market Committee (FOMC) of the Federal Reserve Board (FRB). However, the bandwagon effect was not seen in FOMC.

1.4 Shrinkage and expansion of decision making process

Human decision making is a process that departs from certain premises and heads towards the conclusion. Furthermore, the premise so far is suspected, new premises are explored, heading for the next conclusion (Figure 1). In this way, by repeating the process of shrinkage from the premise towards the conclusion and the process of expansion for new premises, human beings can continue their own decision making and consensus formation with others. Bayesian inference has been used to model human decision-making processes. In Bayesian inference, starting from a premise as a set of hypotheses with likelihood and prior probability, the posterior probability is calculated according to appeared data d. Further, the prior probability of the next step is updated with the posterior probability. The likelihood with the highest prior probability, for example, is the consequence of Bayesian inference, which is the process of shrinkage. The procedure for expansion to explore for new premises is not prepared. On the other hand, in the *inverse Bayesian inference* proposed by Gunji [1], the likelihood itself is updated by using the data history after calculating Bayesian inference. That is, the *inverse Bayesian inference* is the procedure of the extension process to search for new premises.

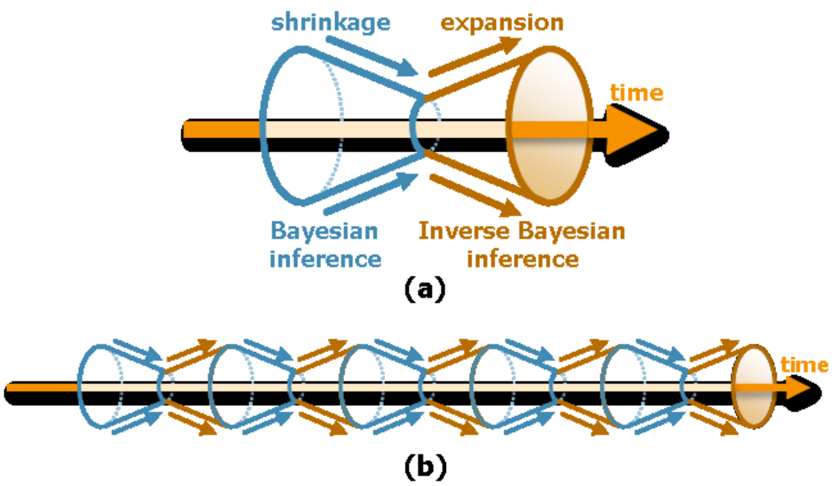


Figure 1 (a) The process of shrinking departing from the premise towards the outcome (Bayesian inference) and the process of expanding to explore new premises (*inverse Bayesian inference*), (b) Continuing decision-making by repetition of shrinkage and expansion process

1.5 Two approaches and three cases to verify

In this thesis, I examine the shrinking and expanding process in the following three cases. That is, "consensus formation aiming for rationality", "decision making process of individuals", and "decision making process of groups" (Figure 2).

The first approach is "symbols and relationships not tied to space". In "consensus formation aiming for rationality", the shrinking process towards the fixation of the logic to solve the problem and the expansion process to fluidization are repeated. In the shrinking process, fixation of logic is attempted while including uncertain elements. Sensitivity analysis of the fixed logic shows the effect of uncertainty overall, and in the expansion process, new elements making up the logic are explored. By using symbols and relationships not tied to space, I show that this expansion process is executed efficiently.

The second approach is "*inverse Bayesian inference*" model. In the "decision making process of individuals", there are a shrinking process to select the optimal solution and an expanding process to search for a different methodology from the optimum solution up to that point. Through the experiments in the maze of virtual space, these shrinking and expanding processes were observed and modeled using *inverse Bayesian inference*. Also in the “decision-making process of groups”, there are aggregation and divergence of opinions by conformity and repulsion with the surroundings. Aggregation of opinions by conforming with the surroundings is a shrinking process, and divergence of opinion by repulsion is an expanding process. These processes were verified by using polls during the election and economic indicators such as stock price and exchange rate and showed that they can be modeled by *inverse Bayesian inference* as well.

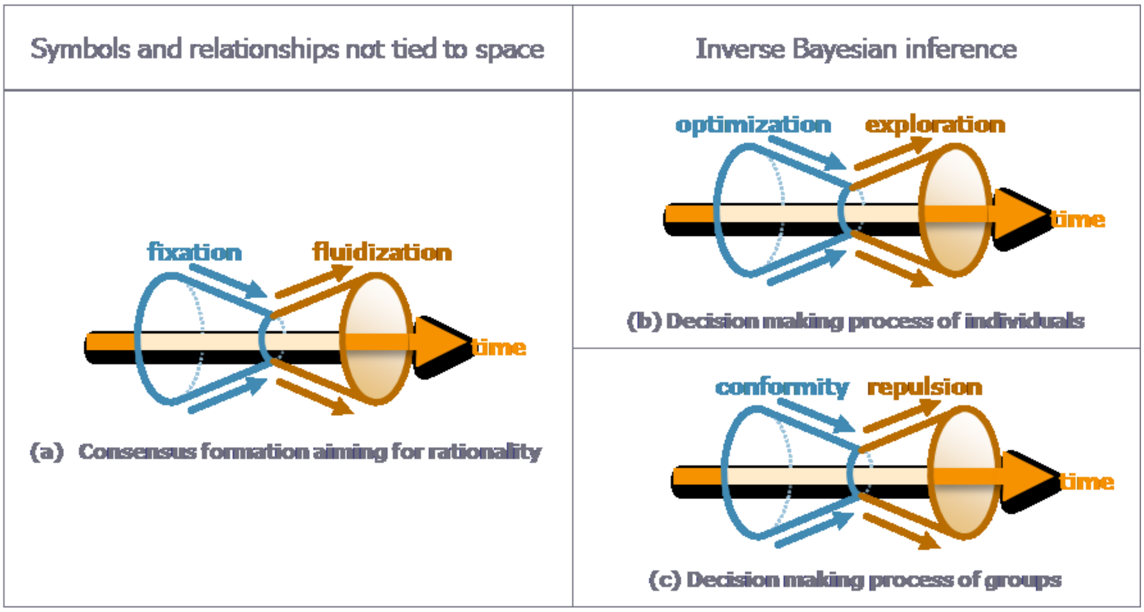


Figure 2 Two approaches and three cases to verify

Chapter 2 Orientation and digest of each chapter

In this chapter, the orientation and the digest of each chapter are described. Figure 3 shows the block diagram of each chapter of this thesis. In **Chapter 2**, I will describe the orientation of my research and the digest of each chapter. In **Chapter 3**, the methodology of *inverse Bayesian inference* is introduced. In **Chapter 4** to **Chapter 6**, the three cases described above are described. And **Chapter 7** concludes and describes the future prospect.

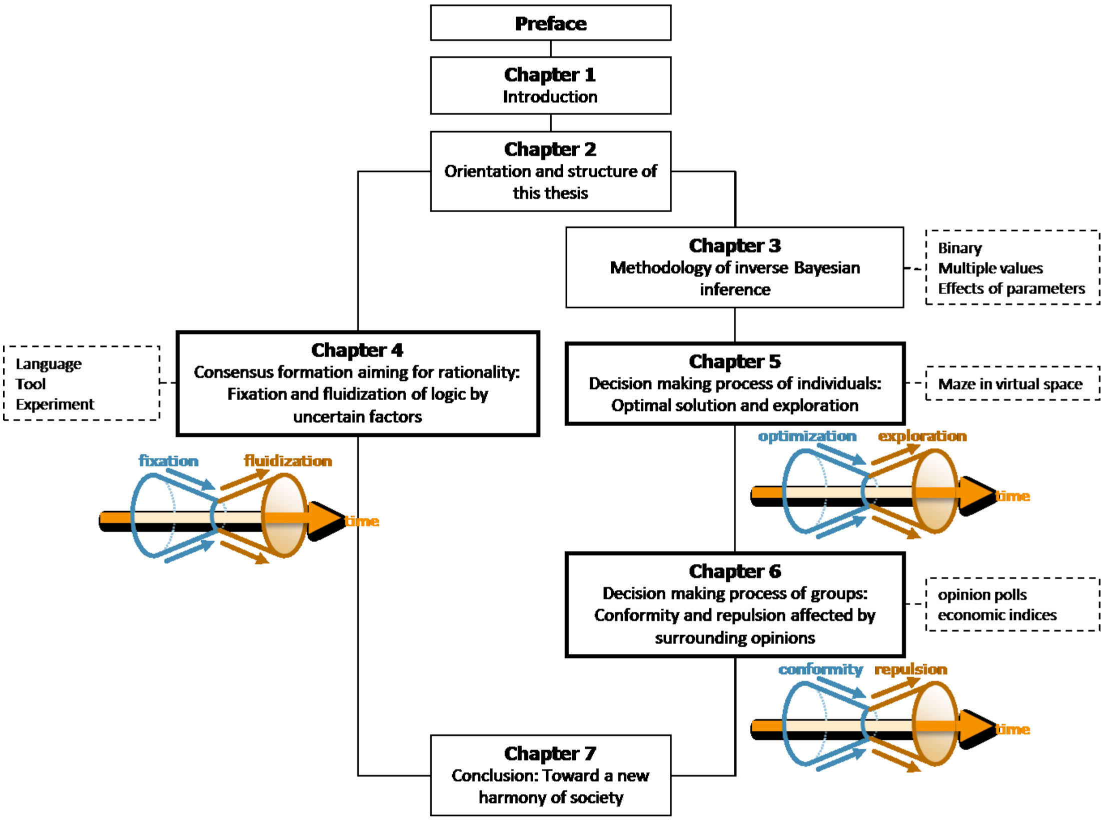


Figure 3 Structure of the chapters in this thesis

In **Chapter 3**, I describe the methodology of the *inverse Bayesian inference* and the influence of the parameters used in the calculation on the inference result. In Bayesian inference, we start on the premise of a set of hypotheses h_i that is likelihood $P^{t}(d|h_i)$and prior probability $P^{t}(h_i)$. The posterior probability $P^{t}(h_i|d)$is calculated from the Bayes theorem by the appearing data d. Further, the prior probability $P^{t+1}(h_i)$ of the next step is updated with the posterior probability $P^{t}(h_i|d)$.

$$P^{t+1}(h_i) = P^t(h_i|d) = P^t(h_i)P^t(d|h_i) / \sum_{k=1..N}(P^t(h_k)P^t(d|h_k))$$

For example, the result of Bayesian inference model is the likelihood with the highest prior probability, and it is the shrinking process. In the *inverse Bayesian inference*, after calculating the Bayesian inference, the likelihood itself is updated using the past history $P^t_{past}(d)$.

$$P^{t+1}(d|h_i) = P^t_{past}(d)$$

In the *inverse Bayesian inference*, there are parameters necessary for calculating the past history. The influence on these estimation results will also be described.

In **Chapter 4**, I describe about the fixation and fluidization of logic in the consensus formation aiming for rationality. For example, when forming consensus on some cost structure, the rationality of each stakeholder is different unless their value criteria are consistent. At least, it is necessary to build logic after sharing same value criteria. For example, let us assume that logic to be shared is defined by elements that can be quantified and relational expressions between elements. Handwriting is sufficient for simple logic, but as the number of elements increases, calculation using a computer is necessary. However, since not all stakeholders have acquired the programming language, spreadsheet software is often used for building and sharing such logic. However, it is reported that there are erroneous inputs in 86% of the spreadsheet, and there is one error per 40 cells. Although the operation of many spreadsheet software is intuitive and the barriers to be learned are low, conventional spreadsheet software has the problems of inducing these erroneous inputs and inhibiting the rational logic fixation and fluidization. Those are, (1) using the same cell for input (= value / formula) and output (= calculation result), (2) using cell address for reference and operand of formula, (3) describing formula with combinations of operators and operands, and (4) only one value that can be handled in each cell. In order to solve these problems, I developed new description language and tools with the following features. (1) clear separation between input and output, (2) using item name (character string) as identifier, (3) describing operator and operand separately, and (4) defining the uncertainty of value. Comparing this language and tools with the conventional spreadsheet software, the working time was reduced by 70% and the erroneous input was reduced by 83%. The logic is smoothly fixed by visualizing the relation between elements and the transition of data, and by analyzing the influence on the whole when the uncertain element changes within the range of possible values, it was possible to smoothly fluidize the logic of searching for new elements.

However, no matter how much attention is paid, humans frequently make unreasonable decisions. In **Chapter 5**, I describe the modeling of individual's decision making process including irrational judgment when going through the maze of virtual space. The subject is instructed to advance as straight forward as possible toward the front of the start position. A total of 8 tasks were performed on the four types of 3D models when only the subjects operate or the software intervenes to change the uncertainty. The distance between the final arrival point and the correct goal, the judgment time, and the ratio of unreasonable choice away from the correct direction did not differ from the difference in uncertainty. However, when modeling with Bayesian inference and *inverse Bayesian inference*, the higher the uncertainty, the more the *inverse Bayesian inference* tends to be able to estimate with higher accuracy than Bayesian inference. Immediately after the start of the experiment, the subject was in the stage of the shrinking process, making decisions that seemed to be the optimal solutions. Subsequently, the subject suddenly noticed that it departed from the goal, and switched to an expanding process to explore for a way to return. In the case of high uncertainty, since such cases frequently occurred, it is considered that the accuracy of the *inverse Bayesian inference* was high. Thus, when the uncertainty is high and the decision-making method itself changes suddenly, it is shown that modeling by *inverse Bayesian inference* is effective.

In **Chapter 6**, modeling decision-making of group will be described. Conforming with surrounding opinions is called a bandwagon effect, and repulsing is called an underdog effect. The bandwagon effect can be said the shrinking process of opinions, and the underdog effect can be said the expansion process of opinion. These effects based on poll results at the time of election have been formulated for a long time and studied through virtual voting experiments. It is also known that the same effect have been recognized in financial markets and enterprises' technology introduction. In this chapter, I examined the transition data of four companies' polls during the 2016 US presidential election. There can be a bandwagon effect if there is a positive correlation between the support rate of the previous survey result and the change of the current support rate, and if there is a negative correlation there is an underdog effect. As a result of the verification, a negative correlation showing an underdog effect was observed. The underdog effect tends to be close battle, in fact, the support rate of both candidates had been around 50% until the end. In addition, according to the support political party, there was a stronger tendency of the underdog effect in the group without the supportive party. The mean squared error between the estimated value by the linear regression model and the actual transition was 1.10% on average for the four companies. In addition, the mean square error of the model by Bayesian inference and *inverse Bayesian inference* are 0.40% and 0.24%, respectively, and the *inverse Bayesian inference* can be said to be the most effective model. This is probably because Bayesian inference has a strong tendency to result in the bandwagon effect, so it was difficult to adapt to this data. On the other hand, *inverse Bayesian inference* is considered to be applicable to both effects. Furthermore, it can also be applied to economic indicators such as stock price and exchange rate, which is another example of group decision making.

Finally, **Chapter 7** describes the concept for new harmony of society and conclusion. In the real world, a large number of decision makers are involved, the uncertainty is high, and various stakeholders make decisions with their respective motives. In order to reasonably form consensus, we have been seeking methodologies to share value criteria and build logic (**Chapter 4**). However, humans frequently make irrational decisions. In order to form an appropriate consensus formation, it was necessary to return to the understanding of the human decision making process itself. The models of individual decision making process including irrational judgment (**Chapter 5**) and the model of group decision process affected by surrounding opinion (**Chapter 6**) will be effective methods. In the future, by incorporating these models, I will establish a methodology for decision making and consensus building that brings new harmony of society.

Chapter 3 Methodology of inverse Bayesian inference

3.1 Case of binary data

3.1.1 Bayesian inference

First, we show a simple example of Bayesian inference. Time series data d:{$d^0..d^{T-1}$(T is data length)} obtained by experiment takes on binary values, that is, 0 or 1. Given N hypotheses, we denote the prior probability of hypothesis h_i ($i = 1..N$) at time step t as $P^t(h_i)$ and the probability (likelihood) that the data at time step t is 1 for a given hypothesis as $P^t(d|h_i)$. Initial values of these probabilities are $P^0(h_i) = 1/N$ and $P^0(d|h_i) = i/(N+1)$. Now, the posterior probability $P^t(h_i|d)$ of each hypothesis h_i according to d^t at each step in this time series of data is successively calculated by equation (1).

$$\begin{aligned} &if\ d^t{=}1 \quad P^t(h_i|d) = P^t(h_i)P^t(d|h_i) \,/\, \textstyle\sum_{k=1..N}(P^t(h_k)P^t(d|h_k)) \\ &if\ d^t{=}\ 0 \quad P^t(h_i|d) = P^t(h_i)(1 - P^t(d|h_i)) \,/\, \textstyle\sum_{k=1..N}(P^t(h_k)(1 - P^t(d|h_k))) \end{aligned} \qquad (1)$$

Next, posterior probability $P^t(h_i|d)$ obtained in this way is substituted in the prior probability of the hypothesis, $P^{t+1}(h_i)$ in the next step.

$$P^{t+1}(h_i) = P^t(h_i|d) \qquad (2)$$

The above process is Bayesian inference. Here, likelihood $P^t(d|h_i)$ of hypothesis h_i for which prior probability $P^t(h_i)$ is maximum shows the change in event probability as inferred from the data.

3.1.2 Inverse Bayesian inference

Next, in *inverse Bayesian inference* [1], the process described above continues with the following process. First, using random number r equal to or greater than 0 and less than $\sum_{k=1..N}(1-P^t(h_k))$, we select hypothesis h_i with the smallest possible prior probability $P^t(h_i)$ satisfying equation (3).

$$Q_{i-1} \le r < Q_i \ where\ Q_i = \textstyle\sum_{k=1..i}(1 - P^t(h_k)), \quad Q_0 = 0 \qquad (3)$$

Likelihood $P^{t+1}(d|h_i)$ of the above hypothesis is now substituted by the most recent moving average $P^{t}_{mov}(d)$ of the frequency of $d^{t}=1$. In other words, this process updates the hypothesis itself. s indicates the interval of moving average.

$$P^{t+1}(d|h_i) = P^{t}_{mov}(d) = \sum_{k=t-s..t}(d^{k} / (s+1)) \quad (4)$$

3.1.3 Features of inverse Bayesian inference

As described above, the equations (1) to (4) are combinations of Bayesian inference and *inverse Bayesian inference*, and therefore are called BIB (both Bayesian and Inverse Bayesian) inference [1]. **Figure 4** shows a simple example of Bayesian inference (a) and BIB inference (b). The binary time series data (1 or -1) used here is randomly generated, and the probability that 1 will occur is 0.85 before 500 steps and 0.25 thereafter (gray). The staircase lines with color indicate the likelihoods with the maximum prior probabilities, the BIB inference follows the sudden change in probability, while Bayesian inference follows asymptotically. In addition, the blue lines show (a) the cumulative average from the start and (b) the moving average with the interval of 25 steps, which change with the same tendency as the result of each inference.

Here, joint probability $P(d, h)$ satisfies equation (5).

$$P(d, h) = P(h|d)P(d) \quad (5)$$

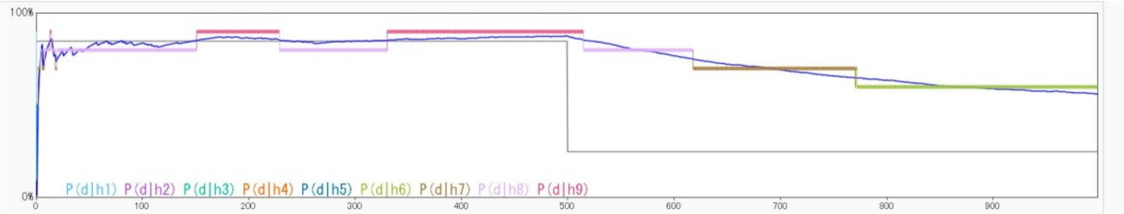
Equation (6) can therefore be obtained from equation (2) of Bayesian inference.

$$P^{t}(d, h) = P^{t+1}(h)P^{t}(d) \quad (6)$$

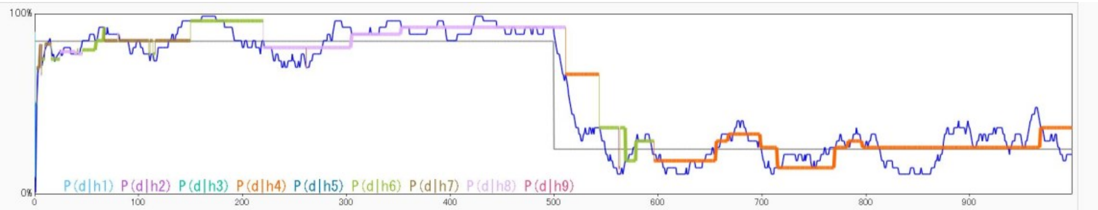
In addition, equation (7) can be obtained from equations (4) and (5) of *inverse Bayesian inference*.

$$P^{t}(d, h) = P^{t}(h)P^{t-1}(d) \quad (7)$$

In equations (6) and (7), the prior probability and the data probability become the temporal converse of each other. In equation (6), the prior probability $P^{t+1}(h)$ is predicted beforehand to directly derive an optimal solution. On the other hand, in equation (7), data probability $P^{t-1}(d)$ is determined by *post-diction* to alleviate previously prescribed rules. That's the reason why the name of this method is called "*inverse" Bayesian inference* [1].



(a)



(b)

Figure 4 (a) Change in cumulative average of the data from the start $P^t_{ac}(d)$ (blue), and likelihood $P^t(d|h_i)$ for maximum prior probability $P^t(h_i)$ by Bayesian inference. (b) Moving average of the data $P^t_{mov}(d)$ whose interval is 25 (blue), and likelihood $P^t(d|h_i)$ by BIB inference. Both data are randomly generated and the probability that 1 will occur is 0.85 when $T \leq 500$, and 0.25 when $500 < T$ (gray).

3.2 Case of multiple data

3.2.1 Bayesian inference

First, we show a calculation method of Bayesian inference. Let $R(t)$:{$R(0)$.. $R(T\text{-}1)$ (T is data length)} be the raw data series, and let d_m:{d_0 .. $d_{M\text{-}1}$ (M is the total number of discrete values)} be the discrete values of the difference between $R(t)$ and $R(t+1)$. The raw data is converted to discrete data sequence $r(t)$:{$r(0)$..$r(T\text{-}1)$ ($r(t) = m$ where $d_m \leq R(t+1) - R(t) < d_{m+1}$)}. Let $P^t(d_m|h_i)$ be the likelihood that hypothesis h_i takes values in each interval m. Given N hypotheses, we denote the prior probability of hypothesis h_i ($i = 1$ .. N) at time step t as $P^t(h_i)$ and the probability (likelihood) that the $r(t) = m$ for a given hypothesis as $P^t(d_m|h_i)$. Initial values of $P^0(h_i)$ are $1/N$, and $P^0(d_m|h_i)$ are normal distributions where their means are $m / M * N$ and their variances are σ^2. Now, the posterior probability $P^t(h_i|d_m)$ of each hypothesis h_i according to $m = r(t)$ at each step in this time series of data is successively calculated by equation (8).

$$P^t(h_i|d_m) = P^t(d_m|h_i)P^t(h_i) \,/\, \Sigma_{k=1..N} P^t(d_m|h_k)P^t(h_k) \quad (8)$$

Next, posterior probability $P^t(h_i|d_m)$ obtained in this way is substituted in the prior probability of the hypothesis, $P^{t+1}(h_i)$ in the next step.

$$P^{t+1}(h_i) = P^t(h_i|d_m) \quad (9)$$

The above process is Bayesian inference. Here, likelihood $P^t(d_m|h_i)$ of hypothesis h_i for which prior probability $P^t(h_i)$ is maximum shows the change in event probability as inferred from the data.

3.2.2 Inverse Bayesian inference

Next, in *inverse Bayesian inference* [1], the process described above continues with the following process. First, using random number r equal to or greater than 0 and less than $\sum_{k=1..N}(1-P^t(h_k))$, we select hypothesis h_i with the smallest possible prior probability $P^t(h_i)$ satisfying equation (10).

$$Q_{i-1} \leq r < Q_i \text{ where } Q_i = \sum_{k=1..i}(1 - P^t(h_k)), \quad Q_0 = 0 \tag{10}$$

Likelihood $P^{t+1}(d_m|h_i)$ of the above hypothesis is now substituted by the most recent moving average $P^t_{mov}(d_m)$ of the frequency of $m = r(t)$. In other words, this process updates the hypothesis itself. s indicates the interval of moving average.

$$P^{t+1}(d_m|h_i) = P^t_{mov}(d_m) = \sum_{k=t-s..t}(r(k) / (s+1)) \tag{11}$$

In order to investigate whether actual data can be modeled by the inference, comparison can be made based on the cumulative value of the likelihood of the hypothesis h^* with the highest prior probability. The model functions obtained by Bayesian inference $M_b(t)$ and BIB inference $M_{bib}(t)$ are as follows:

$$M_b(t+1) = M_b(t) + \sum_{u=0..M-2}(d_u + d_{u+1})\, P^t(d_u|h^*) / 2 \tag{12}$$

$$M_{bib}(t+1) = M_{bib}(t) + \sum_{u=0..M-2}(d_u + d_{u+1})\, P^t(d_u|h^*) / 2 \tag{13}$$

3.3 Effects of parameters

Here, the influences of parameters used in the *inverse Bayesian inference* are described [74]. Figure 5 shows the results of Bayesian inference and *inverse Bayesian inference* of Nikkei Stock Average data. In the following, this data will be used for explanation.

First, with reference to Figure 6, the influence of the random number used in equation (3) will be described. The bar graph of Figure 6 is the distribution of the mean square error when changing the sequence of random numbers (r). The mean square error when choosing the minimum prior probability $P(h_i)$ without using random numbers is 5,143 (one-dotted broken line) and the mean square error when using Bayesian estimation is 7,046 (dotted line). The mean square error between the estimation result and the actual data is distributed lower than Bayesian inference in most cases of BIB inference using random numbers (393 series in 400 series), and the BIB inference result without using random numbers is also lower than the Bayesian inference result.

Next, Figure 7 shows the mean square error between the actual data and the estimated data when the moving average section (*s*) used for BIB inference is changed. Although the error of the BIB estimation is smaller than the error of the Bayesian estimation irrespective of the moving average section s, it shows an irregular behaviour.

Estimation accuracy may be significantly reduced unless the lower limit is given to the prior probability $P(h_i)$. The mean square error when changing the lower limit (*eps*) of the prior probability from 10^{-13} to 10^{-1} is shown in Figure 8. When the lower limit value is 10^{-7} or less, the accuracy of Bayesian inference deteriorates. On the other hand, as the lower limit value increases, the accuracy of BIB inference rises, but there is a tendency to suddenly become worse from around 10^{-2}. In this example, since the number of hypotheses (*N*) is 33, the prior probability is 0.03 if they are equal. If the lower limit value is increased by more than a certain value, it is considered that depending on the number of hypotheses, other prior probabilities are affected and the estimation accuracy worsens.

The influence of the number of hypotheses (*N*) on the estimation result is shown in Figure 9. For both Bayesian and BIB inference, stable accuracy can be obtained with the number of hypotheses greater than a certain value.

Finally, Figure 10 shows the influence of the number of divisions (*M*) when multivalued on the estimation result. While the accuracy of BIB inference is stable, the accuracy of Bayesian inference increases as the number of divisions increases.

As described above, the accuracy of the BIB inference varies depending on the parameters such as the random number sequence, the moving average interval, the lower limit value of the prior probability, the number of hypotheses, and the number of divisions into multilevel values. However, in most cases BIB inference can estimate with higher accuracy than Bayesian inference.

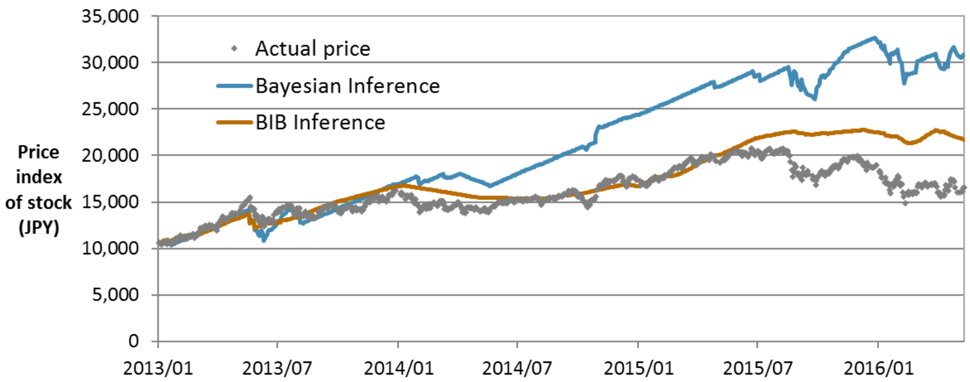


Figure 5 The price index of stock (Nikkei stock average (*T=821*)) and result of Bayesian and BIB inference. Number of hypothesis (*N*) is 33, interval of moving average (*s*) is 32, division number (*M*) is 30, and the limiter of $P(h_i)$ (eps) is 10^{-3} (also in Figure 6 to Figure 10).

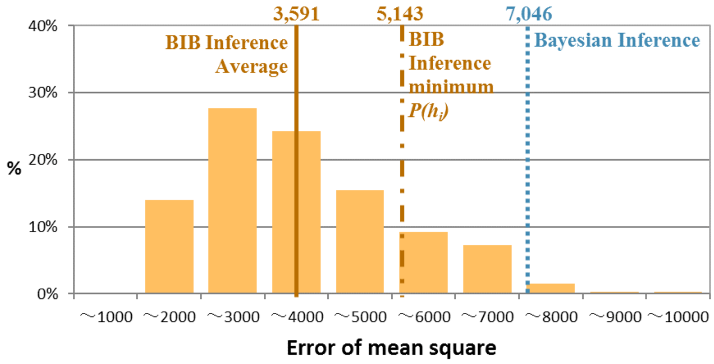


Figure 6 Distribution of mean square error of actual data and inference result when random number sequence used for BIB inference is changed.

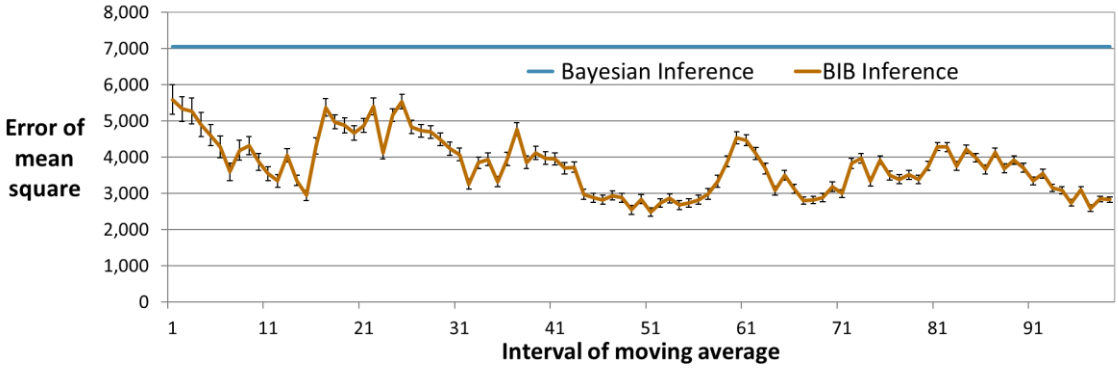


Figure 7 Mean square error when changing the interval of moving average (*S*) used for BIB inference. The graph of the BIB inference result shows the average and standard error using 100 random number series (also in Figure 8 to Figure 10).

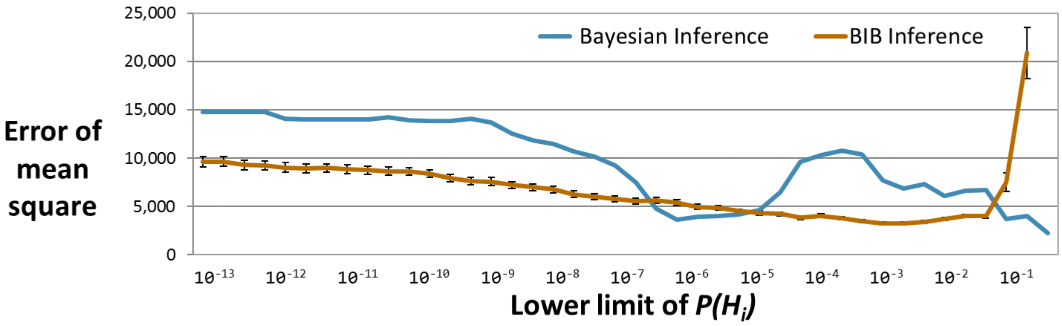


Figure 8 The mean square error when changing the lower limit of the prior probability (*eps*) from 10^{-13} to 10^{-1}.

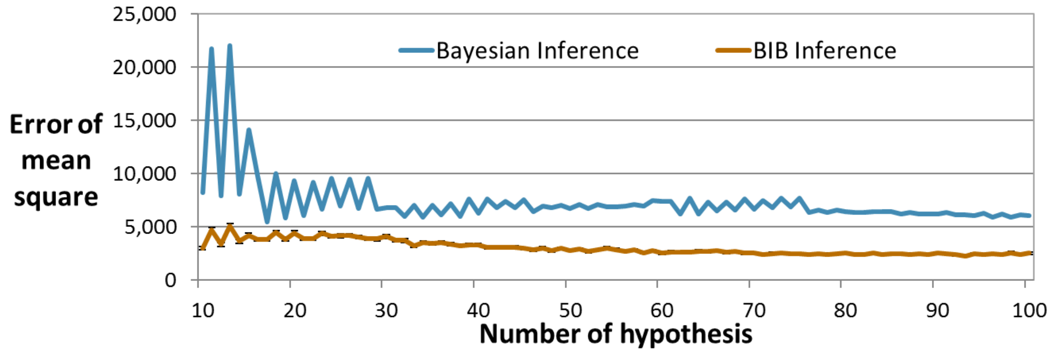


Figure 9 The mean square error when changing the number of hypotheses (*M*) from 10 to 100.

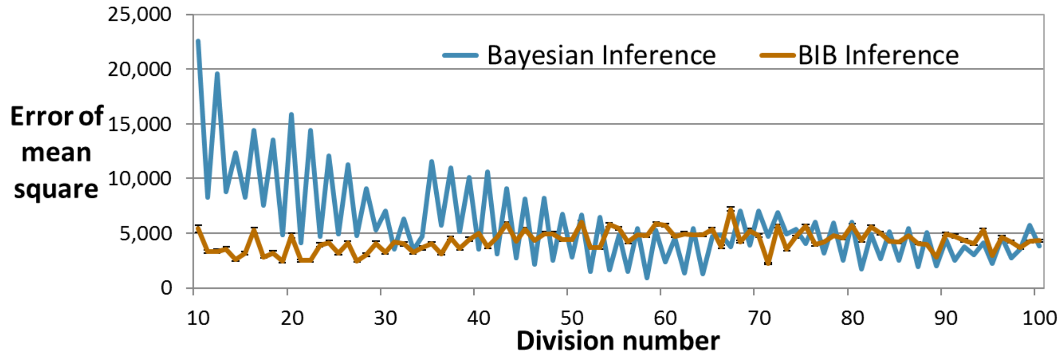


Figure 10 The mean square error when changing the number of divisions (*M*) for multivalued from 10 to 100.

Chapter 4 Consensus formation aiming for rationality: Fixation and fluidization of logic by uncertain factors

Summary of Chapter 4

In this chapter, I describe about the fixation and fluidization of logic in the consensus formation aiming for rationality. For example, when forming consensus on some cost structure, the rationality of each stakeholder is different unless their value criteria are consistent. At least, it is necessary to build logic after sharing same value criteria. For example, let us assume that logic to be shared is defined by elements that can be quantified and relational expressions between elements. Handwriting is sufficient for simple logic, but as the number of elements increases, calculation using a computer is necessary. However, since not all stakeholders have acquired the programming language, spreadsheet software is often used for building and sharing such logic. However, it is reported that there are erroneous inputs in 86% of the spreadsheet, and there is one error per 40 cells. Although the operation of many spreadsheet software is intuitive and the barriers to be learned are low, conventional spreadsheet software has the problems of inducing these erroneous inputs and inhibiting the rational logic fixation and fluidization. Those are, (1) using the same cell for input (= value / formula) and output (= calculation result), (2) using cell address for reference and operand of formula, (3) describing formula with combinations of operators and operands, and (4) only one value that can be handled in each cell. In order to solve these problems, I developed new description language and tools with the following features. (1) clear separation between input and output, (2) using item name (character string) as identifier, (3) describing operator and operand separately, and (4) defining the uncertainty of value. Comparing this language and tools with the conventional spreadsheet software, the working time was reduced by 70% and the erroneous input was reduced by 83%. The logic is smoothly fixed by visualizing the relation between elements and the transition of data, and by analyzing the influence on the whole when the uncertain element changes within the range of possible values, it was possible to smoothly fluidize the logic of searching for new elements. This chapter is based on [75] and [76].

4.1 Fixation and fluidization of logic

There are many scenes where reasonable consensus is formed. In the business scene, stakeholders agree on targets and logic such as “what is appropriate resource allocation?” and “what are the preconditions for achieving the profit target?” For many of the items that make up the logic, quantifiable indicators are used, and relations between symbols indicating items are established. Even if logic is fixed at a certain stage, uncertain items are always included. Also, it is rare that all stakeholders can be convinced with fixed logic. Logic will inevitably be fluidized and relationships between symbols will be rebuilt. That is, fixation of logic is a process of "*Bayesian inference like*" shrinking towards the optimal solution, and fluidization is a process of "*inverse Bayesian inference like*" expansion that searches for other possibilities (Figure 11).

In conventional spreadsheet software, symbols used for logic are fixed as addresses to cells arranged in a lattice pattern (i.e., *space*). Relations and numerical values with other symbols are described in the cell. As soon as focus is removed from the cell, mathematical expressions describing relations between symbols are hidden and only the numerical values of the result are displayed. This interface is effective and intuitive at the stage of logic fixation. However, when fluidizing the logic, there are many disadvantages and it is a cause of erroneous input. In this chapter, therefore, I introduce a methodology to increase the degree of freedom of calculation without fixing symbols and relationships in *space*.

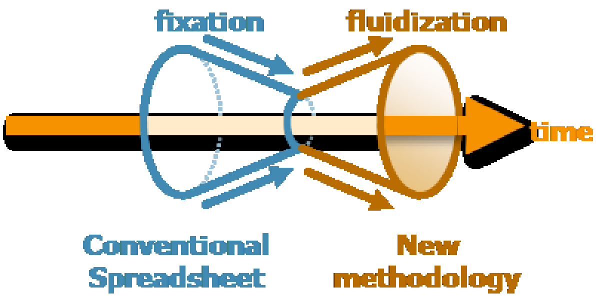


Figure 11 The fixation and fluidization of logic. Conventional spreadsheet software was effective for fixing logic, but it is not suitable for fluidization.

As will be described later, the developed methodology can smooth logic fixation and fluidization and reduce human errors. This methodology has been tried since 2012 and adjusted so that users can freely scrap & build logic. Figure 12 shows the transition of the software reliability growth curve and the number of operators implemented. This methodology is used stably for a period when the curve is flat, and 2013/4 to 2014.4, 2015/10 to 2015/10 can be said to be the period during which this methodology itself has become fluidized.

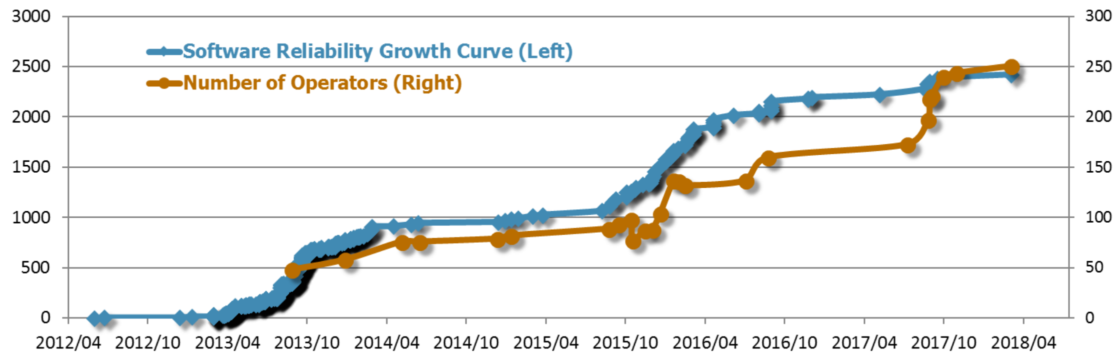


Figure 12 Software reliability growth curve and number of operators implemented.

4.2 Problems in conventional spreadsheets

Spreadsheet software has come to be used in a wide variety of organizations [77] [78] [79], and it is said that about 10 million people create more than a hundred million spreadsheets every year [80]. According to a survey by Gable et al. [81], 84% of users feel that spreadsheet software is extremely important or important and that 96% of those users would recreate a spreadsheet that was somehow lost. In addition, spreadsheets are widely used in mission-critical decisions [82] and more than 88% of users use them in reports to management on the vice-president and higher level [78]. At the same time, the number of basic units or *cells* making up a spreadsheet can range from several thousand to ten thousand [83] [84] and formulas are frequently used. Hall [85] reports that macros are used in 45% of spreadsheets and that spreadsheets are generally becoming more complex in layout as a result of absolute and relative references, logic that includes conditional branching, links with other software programs and databases, etc.

It is also known that spreadsheets include many errors [77] [80] [86] [87] [88] [89] [90]. The results compiled by Panko [80] indicate that errors exist in 86% of spreadsheets and that the percentage of all cells having an error (cell error rate) can run as high as 2.5%. In other words, a large-scale spreadsheet using a large number of cells has a high probability of containing errors. The cause of these errors is not simply carelessness on the part of the spreadsheet creator—it also lies in the limitations of human cognition, a more fundamental problem. It is known that humans will make errors at a rate of about 1/200 when faced with a simple task such as pushing the right switch and at a rate of about 1/50 to 1/20 in more complex work such as programming [80]. Input degree of freedom is particularly high in the case of spreadsheet software where much is left to the discretion of the creator. In work such as this, the occurrence of errors is inevitable.

One specific reason as to why errors occur is that the spreadsheet creator falsely recognizes what type of value a cell indicates. In most spreadsheet software, references and formulas are described using cell addresses. However, as Abraham [91] and Roy [92] have pointed out, it is not a trivial task to automatically extract the value that each cell indicates. In general, the left-most and top-most cells contain row and column descriptions. Consequently, for tables that are exceptionally long in the horizontal or vertical direction, the user may find it necessary at times to scroll back to the top or far left of the spreadsheet or to set the first row or column to be a non-scrolling target. There is also spreadsheet software having a function for assigning names to each cell, but the work of doing so to all cells in a large spreadsheet can be formidable, and since those names will not be reflected when printing out the spreadsheet, such a function is not that effective in actual use. Furthermore, a function for automatically moving relative references when inserting/deleting rows or columns or performing copy-and-paste operations is generally effective, but it may sometimes move a reference contrary to the creator's intention. In the above ways, it can be time consuming to understand the structure of a complicated spreadsheet, particularly one created by another person, which make error discovery all the more difficult. As described later in **Section 4.3** , techniques to help understanding the contents of spreadsheets and methods to reduce errors have been studied. However, while the need for some guidelines in the creation of spreadsheets has been pointed out, there are still fundamental problems inherent to spreadsheet software. One of these is that the same cell is used for both input and output. In many spreadsheet software, a cell in which a computational expression has been input will ordinarily display the result of that computation, although selecting that cell will display the expression. Such specifications are intuitive in nature and have come to be accepted by most users. They are not, however, suitable for obtaining a comprehensive view of computational logic described in a complicated and huge collection of cells. Against this problem, Lotus Improv [93] was an attempt to redefine the way a spreadsheet program should work. It separated the concept of data, views of the data, and formula into three portions. In addition, formulas in conventional spreadsheets are described as a combination of operators and operands. This is a rational method of description for understanding a formula, but when referencing cells of other sheets or files, the resulting character strings can be long hampering legibility.

Another problem is that only one value can be set in each cell. To give calculations a range of fluctuation, that formula needs to be defined using another cell making the spreadsheet all the more complicated. Moreover, when performing sensitivity analysis to investigate the results of varying numerical values in steps or when changing the currency or unit to be displayed, all related cells must be dealt with accordingly. Finally, visualization of calculation results using graphs or other means can be time consuming.

This chapter describes the financial information description language (referred to below as "FDL") and associated visualization and analysis tools that we have developed to solve these problems. A key feature of FDL is clear separations (1) between the input of formulas and numerical values and the output of calculation results and (2) between operators and operands in formulas. Compared to conventional spreadsheet software, the creation time was reduced by about 40% and the number of errors reduced by about 83%.

The remainder of this chapter is organized as follows. **Section 4.3** describes related work especially on domain-specific languages. **Sections 4.4** and **4.5** explain the basic concept and details of FDL. **Section 4.6** introduces visualization and analysis tools using FDL. Experiments comparing FDL and conventional spreadsheet software are described in **Section 4.7** , and **Section 4.8** concludes.

4.3 Some attempts on the problems

The algorithms for extracting spreadsheet errors have come to be researched [94] [95] [96] [97] [98] [99] [100]. For example, Abraham and others [91] [92] [101] [102] infer the data, header, and footer areas of a spreadsheet, infer hierarchical metadata that indicates data content, and extract errors based on compliance with metadata description rules. Relationships of metadata are restricted to trees, preventing DAGs (directed acyclic graphs).

Additionally, to support the understanding of spreadsheet structure, research has been performed on extracting concepts from spreadsheets created by scientists [103] and techniques for extracting metadata for search purposes have been attempted [104] [105] [106]. Other proposals include design methods [107] [108] [109], a cognitive approach [110], a test method [111], and visualization of spreadsheet dataflow [112].

The domain-specific language (DSL) is tailored to a specific application domain [113] [114], and FDL can also be classified as DSL. Recently, compared to general-purpose languages (GPLs), the efficiency and accuracy of DSLs has been investigated.

For example, Kosar et al. [115] investigated the difference in understanding of text-based DSL and GPL programs. In software development, more than 28% of the developer's time is spent on learning and understanding of the program. In their experiments, the participants were asked to analyze the source code of DSL and GPL on the problem of feature diagram, graph descriptions, and graphical user interface, and let them answer the tasks shown on the paper. As a result, they reported that using DSL is more accurate and effective than GPL.

Meliá et al. [116] examined whether DSL notation changes productivity and subjects' satisfaction with graphical or textual differences. Subjects were presented with these two types of DSLs and were asked to discover and improve the errors contained therein. As a result, it became clear that textual notation was more effective. Our FDL is also a textual notation, but as shown in **Section 4.6** , the calculation results are presented in various graphical expressions.

Häser et al. [117] examined whether the language that introduced the concept of the specific domain is more effective than the language not introduced. Languages including business domain concepts were superior both subjectively and objectively than languages consisting only of general keywords. Although FDL is applicable outside the finance area, it also includes functions and concepts suitable for finance calculations.

When experts use marine ecosystem simulation, Johanson et al. [118] compared the effect with GPL (C ++) extended with its own API and DSL composed of concept familiar to experts. Comparative experiments were conducted online for marine experts. Even in this case, DSL is reported to be more effective. The assumed users of FDL are those who actually use spreadsheet software in their business. As described in **Section 4.7** , we recruited them as subjects and conducted experiments to confirm the effect using the developed software.

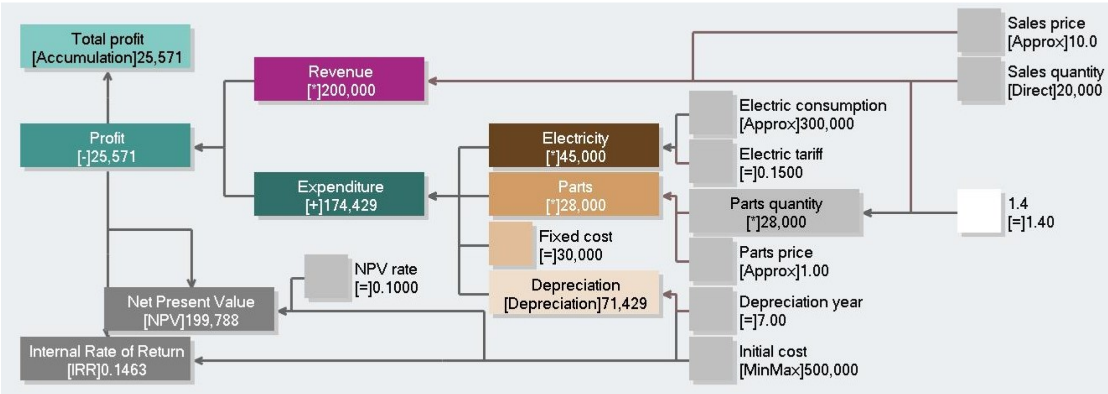
In all of these studies experiments to validate the effect are carefully designed. First, a research goal (**RG**) is defined and it is decomposed into research questions (**RQs**). Next, clarify the independent and dependent variables of the experiment. Dependent variables include objectively measured numerical values and subjective evaluation by questionnaire. Furthermore, a specific hypothesis (**Hi**) is made to clarify the **RQs**. These are organized by Goal, Question, Metrics (GQM) tree, and it is judged from experiment results whether hypotheses are adopted or rejected. Furthermore, threats to validity are verified, and it is confirmed whether the control parameter correctly reflects the hypothesis judgment (internal validity) and whether the hypothesis can be generalized (external validity). This method is also applied to FDL and is described in **Section 4.7** .

4.4 Basic Concept of Financial Information Description Language

4.4.1 Conventional spreadsheet description method

To begin with, we take up a classical example of creating a spreadsheet. In a conventional spreadsheet, item names are usually enumerated along the top-most row or left-most column of the table with time periods or cases enumerated along the other row or column. Other cells describe a certain period (or case) of a certain item. For example, when calculating a firm's state of business for a revenue-and-expenditure structure as shown in **Figure 13(a)**, each item name is entered at the beginning of each row (column A in **Figure 13(b)**) while each period (year, for example) is entered at the beginning of each column (row 1 in **Figure 13(b)**). In addition, rows 5 and 12-19 are items defining numerical values while rows 2-4, 6-11, and 20 are items defining formulas consisting of other items. For example, "Revenue" on row 2 multiplies "Sales price" and "Sales quantity", and cell C2 contains "=C15*C16". In the next period, cell D2 contains "=D15*D16", which can be entered by a function that provides for automatic conversion when performing a copy-and-paste from C2 to D2. However, as mentioned in **Section 4.1** , it is difficult to understand what the addresses referenced by formulas in cells actually indicate with that type of address notation. Furthermore, as the number of items and periods increase, so do the number of cells requiring input. In this example, excluding the first row and first column, numerical values or formulas need to be entered in 215 cells. Consequently, even a simple spreadsheet as in this example holds the possibility of five errors assuming a cell error rate of 2.5%.

Changes in preconditions can also be troublesome in spreadsheet creation. In the spreadsheet of **Figure 13(b)**, for example, changing the value of "Depreciation year" from 7 to another year in the calculation of "Depreciation" would require only that cell B13 be changed since this cell is referenced by the formulas on row 6. On the other hand, changing the fluctuation rate of 1.02 contained in the "Electric Tariff" formulas on row 17 would require that all of those formulas be rewritten. Additionally, when creating graphs, it may be necessary to select multiple cells or, in some cases, to create a separate table.



(a)

	A	B	C	D	E	F	G	H	I
1			1	2	3	4	5	6	7
2	Revenue		=C15*C16	=D15*D16	=E15*E16	=F15*F16	=G15*G16	=H15*H16	=I15*I16
3	Electricity		=C17*C18	=D17*D18	=E17*E18	=F17*F18	=G17*G18	=H17*H18	=I17*I18
4	Parts		=C19*C20	=D19*D20	=E19*E20	=F19*F20	=G19*G20	=H19*H20	=I19*I20
5	Fixed cost		30,000	30,000	30,000	30,000	30,000	30,000	30,000
6	Depreciation		=IF(C1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(D1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(E1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(F1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(G1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(H1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(I1<=$B$13, SLN($B$12, 0,$B$13), 0)
7	Expenditure		=SUM(C3:C6	=SUM(D3:D6	=SUM(E3:E6	=SUM(F3:F6	=SUM(G3:G6	=SUM(H3:H6	=SUM(I3:I6
8	Profit	=-B12	=C2-C7	=D2-D7	=E2-E7	=F2-F7	=G2-G7	=H2-H7	=I2-I7
9	Total profit		=C8	=C9+D8	=D9+E8	=E9+F8	=F9+G8	=G9+H8	=H9+I8
10	Internal Rate of Return	=IRR(B8:Q8							
11	Net Present Value	=NPV(B14,C8:Q8)-$B$12							
12	Initial cost	500,000							
13	Depreciation year	7							
14	NPV rate	10.00%							
15	Sales price		10	10	10	10	10	10	10
16	Sales quantity		20,000	21,000	22,000	24,000	28,000	28,000	28,000
17	Electric Tariff		0.15	=C17*1.02	=D17*1.02	=E17*1.02	=F17*1.02	=G17*1.02	=H17*1.02
18	Electric Consumption		300,000	300,000	300,000	300,000	300,000	300,000	300,000
19	Parts price		1	1	1	1	1	1	1
20	Parts quantity		=C16*1.4	=D16*1.4	=E16*1.4	=F16*1.4	=G16*1.4	=H16*1.4	=I16*1.4

	A	J	K	L	M	N	O	P	Q
1		8	9	10	11	12	13	14	15
2	Revenue	=J15*J16	=K15*K16	=L15*L16	=M15*M16	=N15*N16	=O15*O16	=P15*P16	=Q15*Q16
3	Electricity	=J17*J18	=K17*K18	=L17*L18	=M17*M18	=N17*N18	=O17*O18	=P17*P18	=Q17*Q18
4	Parts	=J19*J20	=K19*K20	=L19*L20	=M19*M20	=N19*N20	=O19*O20	=P19*P20	=Q19*Q20
5	Fixed cost	30,000	30,000	30,000	30,000	30,000	30,000	30,000	30,000
6	Depreciation	=IF(J1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(K1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(L1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(M1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(N1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(O1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(P1<=$B$13, SLN($B$12, 0,$B$13), 0)	=IF(Q1<=$B$13, SLN($B$12, 0,$B$13), 0)
7	Expenditure	=SUM(J3:J6	=SUM(K3:K6	=SUM(L3:L6	=SUM(M3:M6	=SUM(N3:N6	=SUM(O3:O6	=SUM(P3:P6	=SUM(Q3:Q6
8	Profit	=J2-J7	=K2-K7	=L2-L7	=M2-M7	=N2-N7	=O2-O7	=P2-P7	=Q2-Q7
9	Total profit	=I9+J8	=J9+K8	=K9+L8	=L9+M8	=M9+N8	=N9+O8	=O9+P8	=P9+Q8
10	Internal Rate of Return								
11	Net Present Value								
12	Initial cost								
13	Depreciation year								
14	NPV rate								
15	Sales price	10	10	10	10	10	10	10	10
16	Sales quantity	28,000	28,000	28,000	28,000	28,000	28,000	28,000	28,000
17	Electric Tariff	=I17*1.02	=J17*1.02	=K17*1.02	=L17*1.02	=M17*1.02	=N17*1.02	=O17*1.02	=P17*1.02
18	Electric Consumption	300,000	300,000	300,000	300,000	300,000	300,000	300,000	300,000
19	Parts price	1	1	1	1	1	1	1	1
20	Parts quantity	=J16*1.4	=K16*1.4	=L16*1.4	=M16*1.4	=N16*1.4	=O16*1.4	=P16*1.4	=Q16*1.4

(b)

Figure 13 Example of revenue-and-expenditure structure (a) and spreadsheet (b)

4.4.2 Basic Structure of financial information description language

We developed financial information description language (FDL) with the aim of solving the problems shown in **Section 4.4.1** . This language assumes a user on a level capable of inputting numerical values and formulas with spreadsheet software and does not necessarily require skills for high-level programming languages such as C++ and Java. The target of description by FDL is a set of items to be calculated for each period or case such as those shown in **Figure 13(a)**.

A key feature of FDL is that input and output and a formula's operators and operands are clearly separated. In conventional spreadsheets, the same cell is used to perform formula input and output of calculation results. This makes it difficult to understand the relationship between items, which constitutes a fundamental problem with conventional spreadsheets. In contrast, FDL provides for the description of each item's numerical value or formula while the results of calculation are output using a separate tool. Another feature is that the locations for describing the operators and operands of a formula are clearly separated and that all values can be handled as having a range.

Revenue and expenditure is generally hierarchical in structure, but that is not the only possibility. For example, "Initial Cost" in **Figure 13(a)** is used not only in "Depreciation" but in "Internal Rate of Return" too. This constitutes a graph structure having direction but no loops, which is called a directed acyclic graph (DAG). Many general spreadsheet software cannot perform calculations that include loops. The target of FDL is basically DAG as well. Accordingly, to enable this structure to be described in a compact manner while giving it sufficient degree of freedom, we considered it best to list the relationships between items separately. In FDL, each row defines an item by using tabs or commas to delimit the item name, the operator, and operands in that order as shown below. Any characters other than a delimiter may be used to describe the item name and operands, and a text editor or other tools including spreadsheet software may be used to create FDL. The character string itself of each item name may be treated as an identifier and used as an operand of another item.

item	**operator**	operand1	operand2	operand3	…

The operator of an item to be calculated on the basis of other items specifies one of the four basic arithmetic operations, a computational rule such as depreciation or internal rate of return, etc. An operand, meanwhile, specifies another item or numerical value. In addition, the operator of an item specifying a numerical value can specify that standard value or maximum/minimum values, fluctuation rate, period-dependent value, degree of certainty, etc. An item can also be defined in detail using multiple lines.

An example of FDL describing the revenue-and-expenditure structure of **Figure 13(a)** is shown in **Figure 14**. Rows 1-3, 5-10, and 19 are items to be calculated based on other items. "Revenue" multiples "Sales price" and "Sales quantity" using the "*" operator, "Expenditure" sums "Electricity", "Parts", "Fixed Cost", and "Depreciation" using the "+" operator, and "Profit" subtracts "Expenditure" from "Revenue" using the "-" operator.

On the other hand, rows 4 and 11-18 are items specifying numerical values. "Depreciation year", which is used to calculate "Depreciation", specifies the fixed value of 7 years using the "=" operator. "Initial cost", meanwhile, specifies a minimum value of 400,000 and a maximum value of 600,000 using the "**MinMax**" operator. In addition, the "**ValueAndGrowth**" operator for "Electric tariff" specifies an initial value of 0.15 and a fluctuation rate per period of 102%. The "**Direct**" operator used in "Sales quantity" directly sets the values that differ for each period. Finally, the reserved word "**$Setting**" may be set for various types of parameters. On row 20, for example, it is used to specify period length using the "**ProjectPeriod**" operator.

This example, while having an equivalent or greater amount of information than the spreadsheet of **Figure 13(b)**, has no need for describing formulas for all periods and consequently consists of only 59 inputs. The difference in the amount of information between the spreadsheet and FDL lies in the use of value range in the latter. Giving values a range gives the results of calculation a range as well, and such range can be used in sensitivity analysis and elsewhere using the visualization and analysis tools described in **Section 4.6** .

	A	B	C	D	E	F	G
1	Revenue	*	Sales price	Sales quantity			
2	Electricity	*	Electric tariff	Electric consumption			
3	Parts	*	Parts price	Parts quantity			
4	Fixed cost	=	30,000				
5	Depreciation	Depreciation	Initial cost	Depreciation year			
6	Expenditure	+	Electricity	Parts	Fixed cost	Depreciation	
7	Profit	-	Revenue	Expenditure			
8	Total profit	Accumulation	Profit				
9	Internal Rate of Return	IRR	Initial cost	Profit			
10	Net Present Value	NPV	Initial cost	Profit	NPV rate		
11	Initial cost	MinMax	400,000	600,000			
12	Depreciation year	=	7				
13	NPV rate	=	10.00%				
14	Sales price	Approximate	10	90%			
15	Sales quantity	Direct	20,000	21,000	22,000	24,000	28,000
16	Electric tariff	ValueAndGrowth	0.15	102%			
17	Electric consumption	Approximate	300,000	90%			
18	Parts price	Approximate	1.00	90%			
19	Parts quantity	*	Sales quantity	1.40			
20	$Setting	ProjectPeriod	15				

Figure 14 Example of FDL

4.5 FDL Operators

As summarized in **Table 1**, FDL operators consist of data operators that describe the value of an item, function operators that describe the relation between items, user-defined operators, configuration operators that make overall settings, and property operators that make settings related to screen display.

Table 1 Types of FDL operators

Item-related settings	Data operators	Describe item value (**Section 4.5.1**)
	Function operators	Describe relation between items (**Section 4.5.2**)
	User-defined operators	Describe user-defined calculation methods (**Section 4.5.3**)
	Property operators	Make settings related to screen display (**Section 4.5.5**)
Overall settings	Configuration operators	Make overall settings (**Sections 4.5.4** and **4.5.5**)

4.5.1 Data operators describing item value and fluctuation rate

To begin with, we describe data operators that describe item values and their fluctuation rate. In FDL, the user may describe standard value, minimum value, maximum value, and fluctuation rate either directly or indirectly. In **Table 2** below describing these operators, specified operands follow the item name and operator in the order a, b, c, d, … with standard value, minimum value, and maximum value denoted as std, min, and max, respectively.

The data operator "=" in **Table 2** sets a value without any range so that standard value, minimum value, and maximum value are all set to the same value. The "**StandardMinMax**" operator directly sets standard value, minimum value, and maximum value. The "**Approximate**" operator automatically sets minimum value and maximum value from the standard value and degree of certainty. The "**Direct**" operator enables values that change every period to be input directly and those values that do not change from a certain period on to be omitted. In the event of periods that are not filled with data, those denoted with "**…**" take on the nearest specified value and those denoted with ">>>" take on the value interpolated from the previous and subsequent values. The "**Periodic**" operator specifies a value that arises periodically; a time period may be specified to one decimal point. For example, a period specified as 1.5 means a value that arises at the rate of two times per three periods.

Table 2 Data operators describing item values and fluctuation rate

=	Sets a fixed value. a: value, b: unit (may be omitted). std = min = max = a
StandardMinMax	Sets standard value, minimum value, maximum value. std=a, min=b, max=c
MinMax	Sets minimum and maximum values. std=(a+b)/2, min=a, max=b
Approximate	Sets an indeterminate value. std=a, min=a-a*(1-b)/2, max=a+a*(1-b)/2
Direct	Sets value that changes every period (continuation of identical values may be omitted). std = min = max = operands
Periodic	Sets value occurring periodically. a: value, b: period, c: start period (initial period if omitted)
Volatility	Sets standard value and range. a: value, b: range, c: unit (may be omitted). std=a, min=a-b/2, max=a+b/2
ValueAndGrowth	Sets value and fluctuation rate simultaneously. a: value, b: rate with respect to previous period

After setting values of the above type, fluctuation rate, degree of certainty, etc. may be set on other rows. An example is shown in **Table 3**. For example, after setting minimum value and maximum value by "**MinMax**" as described below, the fluctuation rate for every period can be set by the "**Rate**" operator as shown resulting in a standard value that changes in the manner of 150, 180, 195, 225, 225, etc.

item	**MinMax**	100	200		
item	**Rate**	100%	120%	130%	150%

Table 3 Data operators specifying fluctuation rate, certainty, etc.

Growth	Sets a fixed fluctuation rate. a: rate with respect to immediately previous period
Rate	Sets fluctuation rates that change every period. operands: fluctuation rates (continuation of identical values may be omitted)
Certainty	Sets change in certainty of value. operands: certainty (0 – 1.0) for each period

4.5.2 Function operators describing relation between items

Function operators that are used for items calculated from other items calculate basic arithmetic functions, interest, etc. They may also be formulas that combine operators or operators that complement certain types of data. These operators are classified into those that performs calculations every period and those that perform calculations across all periods. Function operators that describe arithmetic functions are listed in **Table 4**. In addition to the four basic arithmetic operations, these include trigonometric functions, permutation and combination functions, and logic functions.

Table 4 Arithmetic function operators

+, -, *, /, ^	addition (a+b+c+d..), subtraction (a-b-c-d..), multiplication (a*b*c*d..), division (a/b/c/d..), raising to a power (a^b)
Free Formula	Formula combining four basic arithmetic operations, raising to a power, arguments, and constants. ex. ((a+b)/100-c^d)*e
Factorial	Factorial. a!
Permutation	Permutation. ${}_aP_b$
Combination	Combination. ${}_aC_b$
sin、 cos、 tan	sine: sin(a), cosine: cos(a), tangent: tan(a)
Pi	pi. No operands.
LOG	Natural logarithm. log(a)
LOG10	log to base 10. log10(a)
Largest	Maximum value. operands: targeted items
Smallest	Minimum value. operands: targeted items
Comparison	Comparison. if a<b then c else d
Comparison0	Comparison with 0. if 0<a then b else c

Sign	Sign. if b<a then 1, else if a<b then -1, else 0. a: value, b: threshold (0 if omitted)
Accumulation	Accumulated value up to that period. a: targeted item
Range	Difference between maximum value and minimum value. a: targeted item
Delay	Reference past value. a: target item, b: delay period (1 if omitted)
Difference	Difference with past value. a: targeted item, b: delay period (1 if omitted)
Multiple	Set items separately for each period. operands: item names for each period; continuation of identical item names may be omitted
FirstPeriod, $LastPeriod	Calculate using the value of the immediately previous period. Set the value of the initial period with FirstPeriod and reference the value of the immediately previous period with $LastPeriod on another row.
GetProjectPeriod	Get length of project period. No operands.
GetNthPeriod	Get value of Nth period of another item. a: item, b: period
GetPeriod	Get value of existing period. No operands.
SumProduct	Multiply two items every period and get sum total. a: item 1, b: item 2, c: period (all periods if omitted)
AverageSumProduct	Multiply two items every period and take average. a: item 1, b: item 2, c: period (all periods if omitted)
:=	Define item of same value by a different name.
Case	Specify items for each case. operands: item names or numerical values for each case
SumCase	Sum values for all cases. a: target items
Polyline	Calculate target value by interpolation. a: X-axis item name, b: X-axis data item name, c: Y-axis data item name
Stepwise	Calculate target value in a stepwise manner. a: X-axis item name, b: X-axis data item name, c: Y-axis data item name

In **Table 4**, Free Formula is a formula that describes any combination of operators and operands. It describes operands in the order of a, b, c, ... as shown below.

item	(a+b*100)/c	itemA	itemB	itemC

Here, the formula uses a, b, c, ... that indirectly represents operands, but this hinders understanding of the actual items referenced, which corresponds to the fundamental problem discussed in **Section 4.4.1** . However, it is also possible to directly enter (“itemA” + “itemB” * 100)/“itemC” to avoid use of such an indirect expression. On the other hand, operands with long character strings can hinder legibility making it necessary to use a new delimiter to set off item names, but this would then limit the characters that can be used in item names. In FDL, preference is placed on separating operators and operands.

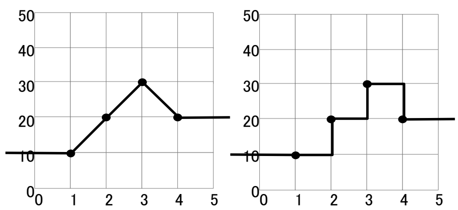
The “Multiple” operator can be used when computational logic changes every period. Its operands consist of separately defined computational-logic items enumerated for each period. Additionally, when referencing the value of the immediately previous period as in row 17 of **Figure 13(b)**, the “FirstPeriod” operator defines the initial period while the second and later periods are defined on separate rows. At this time, the reserved word “$LastPeriod” can be used to reference the value of the immediately previous period.

Furthermore, when attempting to combine data using a formula, the relationship between that data may be too difficult or complicated to express. Consequently, while data having some kind of relationship may exist, it may not be possible to express that relationship by some sort of approximation formula. For example, the data relationship shown in Figure 15(a) can be described as follows using the “Polyline” operator. Here, the value Y1 is 25 achieved by interpolation.

Y1	Polyline	X	dataX	dataY	
dataX	Direct	1	2	3	4
dataY	Direct	10	20	30	20
X	=	2.5			

Similarly, the stepwise-changing data relationship shown in **Figure 15(b)** can be described as follows using the “Stepwise” operator. Here, Y2 turns out to be 20.

Y2	Stepwise	X	dataX	dataY	
dataX	Direct	1	2	3	4
dataY	Direct	10	20	30	20
X	=	2.5			



(a) (b)

Figure 15 Example of relationships difficult to express as a formula

Interest, depreciation, tax, and other function operators commonly installed in spreadsheet software are listed in **Table 5**.

The function operators so far described are used in items that perform calculations every period. In contrast, operators that perform calculations across all periods such as internal rate of return (IRR) and net present value (NPV) are listed in **Table 6**.

Additionally, using certain operators on multiple rows, it is possible to perform processing to limit the results of an item defined by arithmetic functions so as not to exceed a threshold. These operators are listed in **Table 7**.

Table 5 Interest, depreciation, tax, and other function operators

Payment	Total payment for constant periodic payments. a: interest, b: payment period, c: lending amount
InterestPayment	Interest payment for constant periodic payments. a: interest, b: payment period, c: lending amount
PrincipalPayment	Principal payment for constant periodic payments. a: interest, b: payment period, c: lending amount
InterestPaymentEP	Interest payment for fixed rate loan. a: interest, b: payment period, c: lending amount
TotalInterestPayment	Interest payment for constant periodic payments when borrowing every period. a: interest, b: payment period, c: lending amount
TotalPrincipalPayment	Principal payment for constant periodic payments when borrowing every period. a: interest, b: payment period, c: lending amount
Depreciation	Straight-line depreciation allowance. a: estimated value, b: depreciation period, c: start period
DepreciationFD	Declining-balance depreciation allowance. a: estimated value, b: remaining value, c: depreciation period, d: start period (initial period if omitted)
TotalDepreciation	Depreciation allowance for assets arising every period. a: estimated value, b: depreciation schedule
RealEstateTax	Fixed assets tax. a: acquisition price, b: depreciation period, c: tax rate

Table 6 Function operators that perform calculations across all periods

TotalAmount	Cumulative total of all periods. a: target item, b: target period (all periods if omitted)
Average	Average of all periods. a: target item, b: target period (project period if omitted)
IRR	Internal rate of return. a: investment amount, b: recovered item, c: number of calculation years (project period if omitted)
NPV	Net present value. a: investment amount, b: recovered item, c: discount rate
PaybackPeriod	Period that initially has a positive value (recovery period). a: target item

Table 7 Function operators performing post-calculation processing

LimitMinimum	Lower limit. a: lower limit
LimitMaximum	Upper limit. b: upper limit
LimitMinMax	Lower and upper limits. a: lower limit, b: upper limit
LimitMinimum0	Always a positive value. No operands.
Integer	Round down to integer. No operands.
RoundUp	Round up to integer. No operands.

4.5.3 User-defined operators for extending logic

The user can register self-extended computational logic as user-defined operators. To begin with, the reserved word "$CustomBegin" is entered on the first row followed by the name of the customized operator and a list of operands. Then, using these operands, the computational logic of this customized operator is described on multiple rows. Finally, "$CustomEnd" is entered on a new row followed by a formula describing the calculation to perform. The following defines the operator "StockSalesRatio" that calculates the inventory rate for every period from the change in quantity purchased and quantity sold that fluctuates every period. The application of "StockSalesRatio" is then shown for three cases.

$CustomBegin	StockSalesRatio	Quantity purchased	Quantity sold
Total quantity purchased	Accumulation	Quantity purchased	
Total quantity sold	Accumulation	Quantity sold	
Stock on-hand	-	Total quantity purchased	Total quantity sold
$CustomEnd	a/b*100	Stock on-hand	Quantity sold

Inventory rate A	StockSalesRatio	Quantity purchased A	Quantity sold A
Inventory rate B	StockSalesRatio	Quantity purchased B	Quantity sold B
Inventory rate C	StockSalesRatio	Quantity purchased C	Quantity sold C

4.5.4 Configuration operators for describing various settings

Configuration operators are used for making overall settings such as project name and currency-related information independent of particular items. A configuration operator follows "$Setting," which is entered at the beginning of a row in place of an item name. These configuration operators are listed in **Table 8**.

Table 8 Configuration operators describing various types of settings

ProjectName	a: project name
ProjectPeriod	a: project period
PeriodAsVal	Set value of period. a: increment, b: initial value, c: unit
PeriodAsAnnual	Set value of period as western year, etc. a: initial value, b: unit
PeriodAsMonth	Set value of period as month. a: initial value, b: unit
PeriodAsHour	Set value of period as hour. a: initial value, b: unit
Currency	a: name of basic currency
Currency2	Name of secondary currency. a: currency name, b: converted value from basic

	currency unit, c: if entered, b/c is converted value
Currency3	Name of tertiary currency. a: currency name, , b: converted value from basic currency unit, c: if entered, b/c is converted value
CaseName	Set case name. operands: name of each case
Include	Read in another sheet. a: sheet name

4.5.5 Property and configuration operators related to screen display

The results of calculating the items described in FDL are displayed by the visualization and analysis tools described in the following section. We here describe the property operators that make settings related to the display method used at that time. In **Table 9**, the "Color" and "FontColor" operators set the background color and font color when displaying individual items. The colors set in spreadsheet software may be used here.

Table 10 lists configuration operators related to the entire screen display. They can be used to specify number of digits, currency unit, etc. to be displayed. In the table, "PeriodMessage" and "FloatingMessage" are operators for adding comments to a graph, etc. "GroupName" is an operator for assigning a name to a group of items and setting the level of that group if hierarchical. Describing the same property operator for multiple items increases the number of FDL rows thereby degrading readability. For this reason, an operator not requiring operands may be specified for multiple items simultaneously in the following way.

$Setting	Integer	itemA	itemB	itemC	itemD

Conversely, "SetOperators" may be used to simultaneously specify multiple operators for a single item as shown below.

item	SetOperators	LimitMinimum0	Integer	ShowAsBar	KeepDigit

In addition, a range of text enclosed by "/*" and "*/" and a row containing "//" are treated as comments and excluded from read-in.

Table 9 Item-oriented property operators

Color	Set item background color. a: red, b: green, c: blue (0..255)
FontColor	Set item font color. a: red, b: green, c: blue (0..255)
Unitname	Set unit. a: unit name
KeepDigit	Leave unchanged at time of digit/currency conversion for other than monetary amounts. No operands.

Better	Set item for which increase is desirable. Item to be increased when moving slider at bottom of screen. No operands.
NoNeedTotal	Set item requiring no total value at far right of financial index table. No operands.
Show0thPeriod	Display only 0th period on financial index table. a: if another item name is set, display on that row.
Hide	Disable display of item name. No operands.
ShowHereAgain	Reshow previously appearing item. No operands
UnCheckItem	Disable display on graph. No operands.
ShowAsBar	Display as bar graph. No operands.
SetOperators	Set operators. operands: operator names

Table 10 Configuration operators related to screen display

DigitNormal	Display all digits. No operands.
DigitK, DigitM, DigitG	Display digits in terms of K(10^3), M(10^6), and G(10^9). No operands.
SetCurrency1	Display in basic currency. No operands.
SetCurrency2	Display in secondary currency. No operands.
SetCurrency3	Display in tertiary currency. No operands.
StandardValue MinimumValue MaximumValue	Switch between display of standard value, minimum value, and maximum value. No operands
GroupName	Display item group name in table. a: group name, b: hierarchical level
PeriodMessage	Specify a message for every period displayed on a graph, etc. operands: message for every period
PeriodMessageEach	Specify a message for a period displayed on a graph, etc. a: period specifying message, b. message
FloatingMessage	Floating message. operands: message
MessageAttribute	Floating message attributes. a: font color (red), b: font color (green), c: font color (blue), d: background color (red), e: background color (green), f: background color (blue), g: background opacity, integer in range of 0..255 for above attributes, h: font size, i: center text (=1) or align text left (=0)
CheckOnly	Set items to be displayed on graph. Unchecks all items other than those entered as operands.
CaseSelect	Set display/non-display for each case. operands: display (=1) or non-display (=0)
BackgroundFile	Specify file name of background image. a: file name of background image
FontName	Specify font name. a: font name

4.5.6 Compatibility with GPL

FDL is designed from scratch without any relationship to an existing language (i.e. language invention pattern [113]). Even the way of description of the operators and operands are different between FDL and the conventional spreadsheets, their semantics are equivalent. The FDL format (ex. **Figure 14**) can be automatically converted to the conventional spreadsheet format (ex. **Figure 13(b)**), although some information such as value range would have to be omitted. Conversely, a conventional spreadsheet, which may describe item names in a hierarchical manner, can nevertheless be converted to FDL by combining the techniques of Abraham [91] and Roy [92] with the FDL “GroupName” operator and item names.

Furthermore, FDL can be converted to the general purpose programing language (GPL) such as C++. **Figure 16** is an example of automatically converted code from FDL (Figure 14), by replacing item’s name into proper variable names (from space characters to under-bar characters, etc.), and each code is exported recursively using the relationship between the items. In this sense, FDL is a DSL which is open to the GPL, and can be reuse for various purposes.

```
class FDL2CppClass
{
public:
#define ProjectPeriod   15
    double  _RATIO;
    double  Revenue[ProjectPeriod], Electricity[ProjectPeriod], Parts[ProjectPeriod], Fixed_cost[ProjectPeriod];
    double  Depreciation[ProjectPeriod], Expenditure[ProjectPeriod], Profit[ProjectPeriod], Total_profit[ProjectPeriod];
    double  Internal_Rate_of_Return, Net_Present_Value, Initial_cost[ProjectPeriod], Depreciation_year[ProjectPeriod];
    double  NPV_rate[ProjectPeriod], Sales_price[ProjectPeriod], Sales_quantity[ProjectPeriod], Electric_tariff[ProjectPeriod];
    double  Electric_consumption[ProjectPeriod], Parts_price[ProjectPeriod], Parts_quantity[ProjectPeriod];

    FDL2CppClass()
    {
        _RATIO = 0.5;
    }
    void    calculation()
    {
        for (int t = 0; t < ProjectPeriod; t++)
        {
            Sales_price[t] = (1.0 - (1.0 - 0.9000)/2) * 10.0 * (1.0 - _RATIO) + (1.0 + (1.0 - 0.9000)/2) * 10.0 * _RATIO;
            if (0 == t) Sales_quantity[t] = 20000;
            else if (1 == t) Sales_quantity[t] = 21000;
            else if (2 == t) Sales_quantity[t] = 22000;
            else if (3 == t) Sales_quantity[t] = 24000;
            else Sales_quantity[t] = 28000;
            Revenue[t] = Sales_price[t] * Sales_quantity[t];
            Electric_tariff[t] = 0.1500 * pow(1.02, t);
            Electric_consumption[t] = (1.0 - (1.0 - 0.9000)/2) * 300000 * (1.0 - _RATIO) + (1.0 + (1.0 - 0.9000)/2) * 300000 * _RATIO;
            Electricity[t] = Electric_tariff[t] * Electric_consumption[t];
            Parts_price[t] = (1.0 - (1.0 - 0.9000)/2) * 1.00 * (1.0 - _RATIO) + (1.0 + (1.0 - 0.9000)/2) * 1.00 * _RATIO;
            Parts_quantity[t] = Sales_quantity[t] * 1.4;
            Parts[t] = Parts_price[t] * Parts_quantity[t];
            Fixed_cost[t] = 30000;
            Initial_cost[t] = 400000 * (1.0 - _RATIO) + 600000 * _RATIO;
            Depreciation_year[t] = 7.00;
            Depreciation[t] = _Depreciation(Initial_cost[t], Depreciation_year[t]);
            Expenditure[t] = Electricity[t] + Parts[t] + Fixed_cost[t] + Depreciation[t];
            Profit[t] = Revenue[t] - Expenditure[t];
            if (0 == t) Total_profit[t] = Profit[t]; else Total_profit[t] = Total_profit[t - 1] + Profit[t];
            NPV_rate[t] = 0.1000;
        }
        Internal_Rate_of_Return = _IRR(Initial_cost, Profit);
        Net_Present_Value = _NPV(Initial_cost, Profit, NPV_rate);
    }
};
```

Figure 16 Example of C ++ code automatically generated from FDL

4.6 FDL Visualization and Analysis Tools

The following describes the tools for reading FDL-described content and visualizing calculation results and for performing a variety of analyses.

4.6.1 Basic calculation flow

The visualization/analysis tools read the specified FDL file and store its content in memory. At this time, any detected errors, such as when an item used as an operand is not defined or when an undefined operator is used, will be displayed as a warning together with their corresponding row numbers. Since FDL assumes a DAG structure, following the operands of each item will necessarily lead to terminating items for which values have been set. Accordingly, the values of all items can be obtained by creating a graph for which calculation of all items may or may not be completed and performing recursive calculations while following operand-acting items. For an item whose value has already been set, its standard value, minimum value, and maximum value will be calculated using the set value, fluctuation value, etc. On the other hand, for an item calculated from other items, the maximum and minimum values of those operands will be combined so that the maximum and minimum values of the calculation results will take on appropriate values. For example, denoting the minimum and maximum values of operands a and b as a_{min}, b_{min} and a_{max}, b_{max}, respectively, the minimum and maximum values of the sum of those operands will be calculated as $a_{min}+b_{min}$, $a_{max}+b_{max}$ while that of the difference of those operands will be calculated as $a_{min}- b_{max}$, $a_{max}- b_{min}$.

4.6.2 Visualization functions

This section provides an overview of the FDL visualization tools. Although the display method for each screen can be varied by tool-specific operations, it can also be set beforehand by FDL.

An example of displaying the initial fiscal year (initial period) for the screen showing the relationships among items is shown in **Figure 17**. For each item, the name is displayed in the upper row and the operator and the value are displayed in the lower row. In this screenshot, the mouse cursor is hovering over the item “Total profit”. This is the cumulative value of “Profit”, which is calculated from “Revenue” and “Expenditure”. In other words, this screen enumerates hierarchically related items and displays calculation results and operators. In addition, the bars displayed at the top of the screen show the minimum values, maximum values, and standard values of selected items. In the case of items defined by data operators, performing a drag operation on those bars changes those values enabling the results of recalculations to be reflected in real time. In addition to such a tree structure, it can be displayed with a network structure connecting related items. **Figure 13(a)** in **Section 4.4.1** is another example of this network display like UML (Unified Modeling Language) diagram, and the position of each item is specified by the user.

An example of a screen displaying calculation results in table format is shown in **Figure 18**. Here, calculation results can be switched among standard value, minimum value, and maximum value. In addition, "Internal Rate of Return" and "Net Present Value" are examples of items calculated across all periods and displayed at the left-most column in the table. Other items can be displayed in a similar manner using the "**Show0thPeriod**" operator. These results can be output to a spreadsheet with formulas included similar to the one shown in **Figure 13(b)**. In addition, the value-setting window at the lower right in **Figure 18** can be used to change the value of the selected item continuously by a drag operation thereby recalculating results in real time. This value-setting window can be effectively used on all of the screens shown below.

Although FDL can handle minimum and maximum values, these values may not be the worst and best values. For example, if "Parts price" decreases while "Sales price" increases, "Profit" improves, which is a favorable outcome. If the "**Better**" operator in **Table 9** is set for an item for which an increase is desirable, the slider at the bottom left of the screen can be operated to continuously simulate conditions from worst to best.

An example of a screen displaying graphs that depict the change in value of various items is shown in **Figure 19**. The vertical bars on a line graph indicate the range that that item can take on for that period from its minimum value to maximum value. The items to be displayed here can be selected using the checkboxes in the item list on the right, but the "**CheckOnly**" operator may also be used to set them beforehand. In addition, the "**ShowAsBar**" operator may be used to display specified items as bar graphs.

In addition to the above, calculation results may also be visualized in the form of bubble charts or pie charts for any items as shown in **Figure 20** and **Figure 21**, respectively. For a pie chart, the breakdown of an item specified with the operator "+" is displayed in a hierarchical manner.

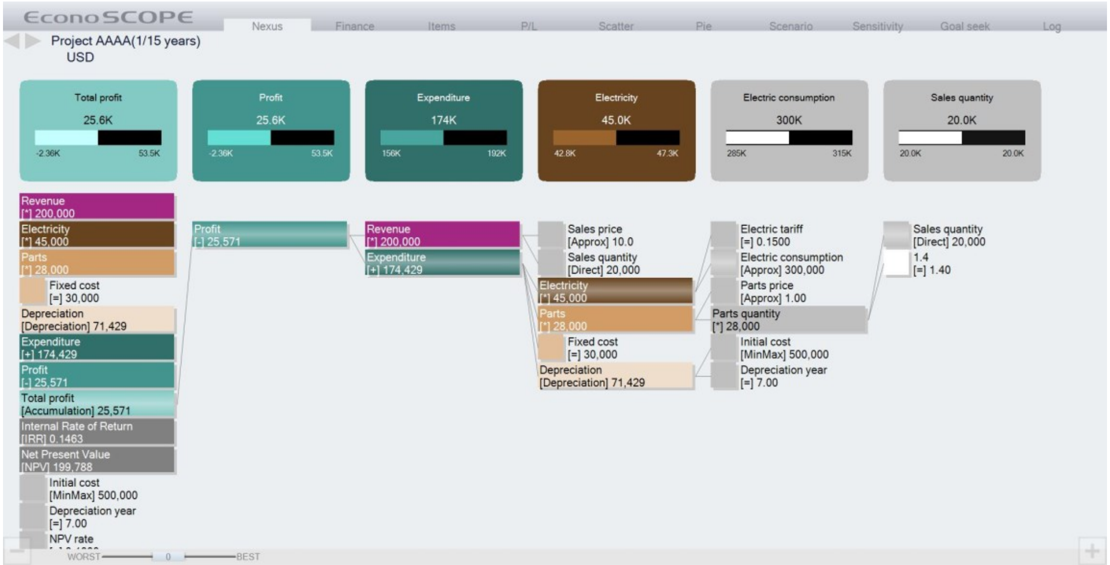


Figure 17 Screenshot of inter-item relationships

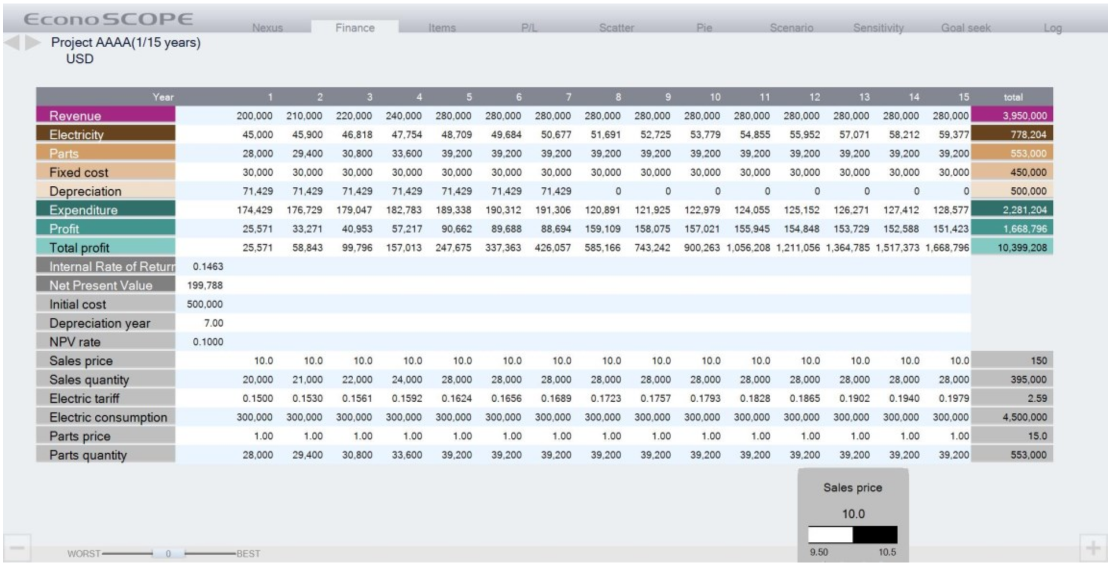


Figure 18 Screenshot of calculation results

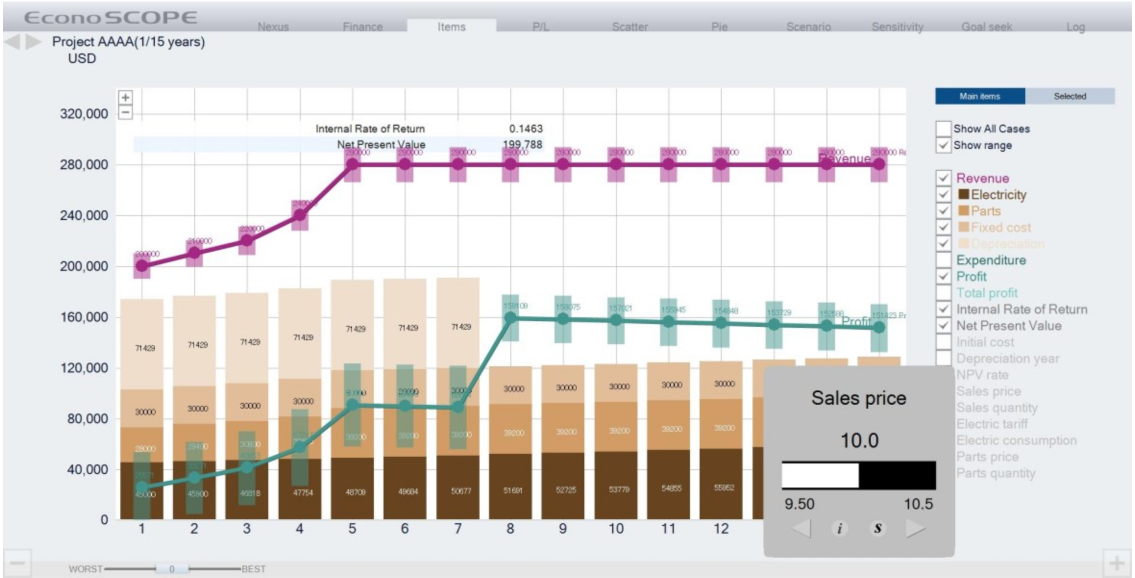


Figure 19 Screenshot of graphs of calculation results

Figure 20 Screenshot of bubble chart

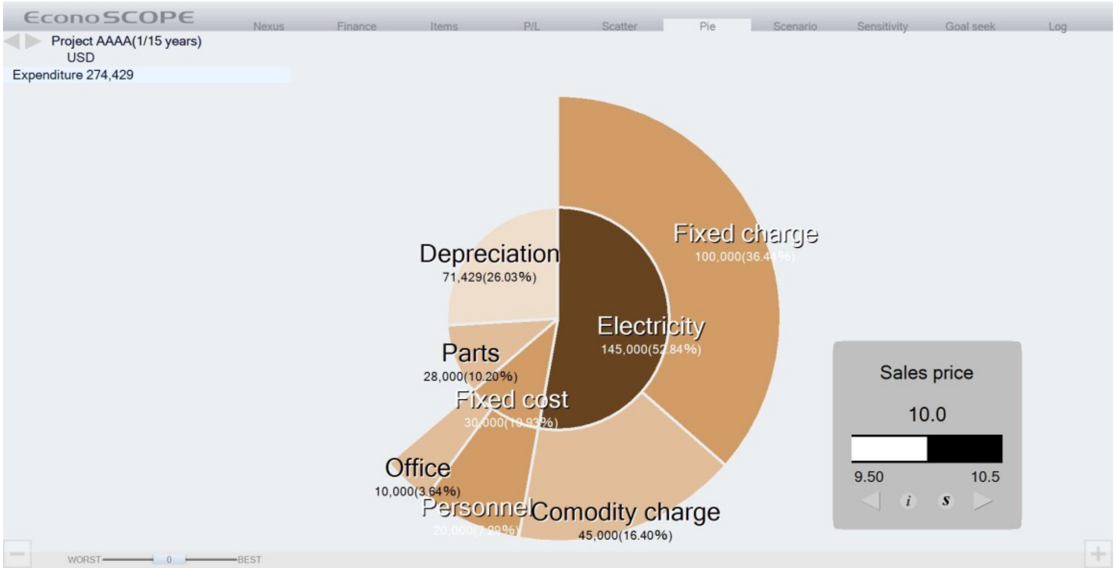


Figure 21 Screenshot of pie chart

4.6.3 Analysis functions

To assess risk factors and simulate diverse scenarios, it is essential to quantitatively calculate how the target index changes when varying preconditions, or alternatively, what preconditions must be satisfied to reach the target. The former process uses sensitivity analysis while the latter uses Goal Seek analysis. These analysis methods have been troublesome to perform in conventional spreadsheets, but FDL visualization and analysis tools can execute them by simply selecting desired items.

A screenshot of a spider chart in sensitivity analysis is shown in **Figure 22**. A spider chart is a graph that shows change in the index item corresponding to change in element items. After selecting the items to be graphed, the spider chart displays the results of calculating the index item for combined conditions consisting of equally divided values from the minimum value to maximum value of each element item. In the following example, the spider chart shows calculation results for “Internal Rate of Return”, the index item, on the vertical axis when varying the values of the element items “Sales price” on the horizontal axis and “Initial cost” on each line graph.

Next, a screenshot of a tornado chart in sensitivity analysis is shown in **Figure 23**. A tornado chart is a graph that shows how change in each element affects the index arranged in descending order of effect. Here, the index item is calculated for equally divided values of each element from its minimum value to maximum value and the minimum and maximum values of the resulting calculations are displayed for each item. The change rates of elements are frequently calculated using the same value. In this case, the “Approximate” operator may be used to set the same degree of certainty. However, if comparing with the same change rate is unreasonable, either degree of certainty or minimum/maximum values may be set as necessary. The symbol “◊” or “♦” affixed to each item in the chart indicates a positive or negative correlation, respectively, between the element and index.

Finally, a screenshot of the Goal Seek function is shown in **Figure 24**. On selecting the index item and inputting a target value, this function back-calculates the values of all element items defined by data operators so that the index item becomes the target value. In some cases, the index item may never arrive at the target value. In the following example, the index item “Internal Rate of Return” will achieve the target value of 0.3 if the element item “Initial cost” can be decreased from 500,000 to 291,455. Here, however, the values calculated for element items such as “Depreciation year” and “Fixed cost” are unrealistic while “NPV rate” has absolutely no effect on the index reaching the target value. It can therefore be seen that the target cannot be achieved solely on the basis of these items. To prevent a search with such unrealistic values, the “LimitMinMax” operator can be used to limit the values that an item can take on. It is also possible to calculate combinations for achieving the target value while simultaneously varying multiple element items.

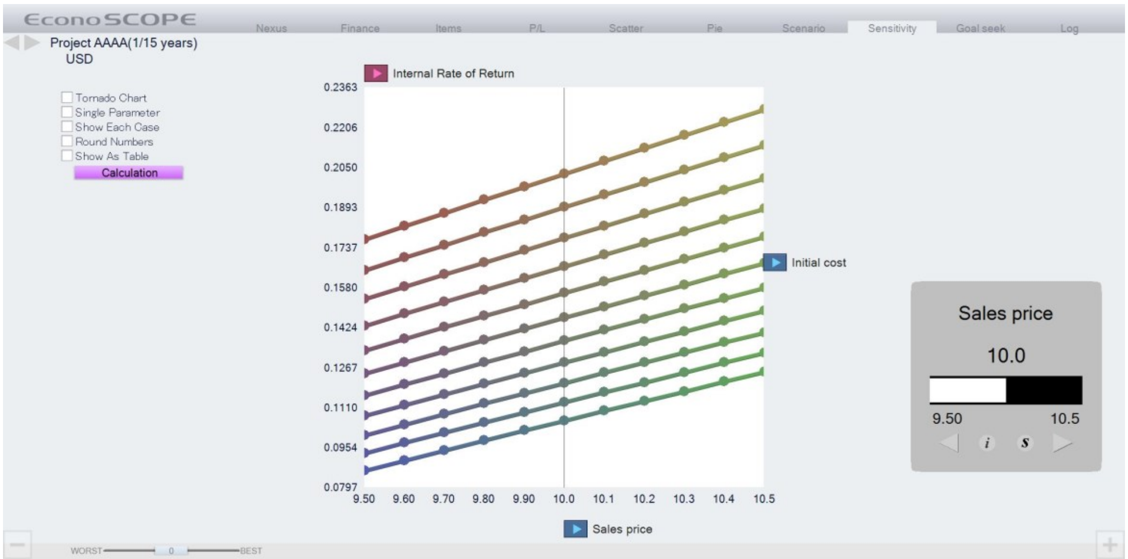


Figure 22 Screenshot of spider chart

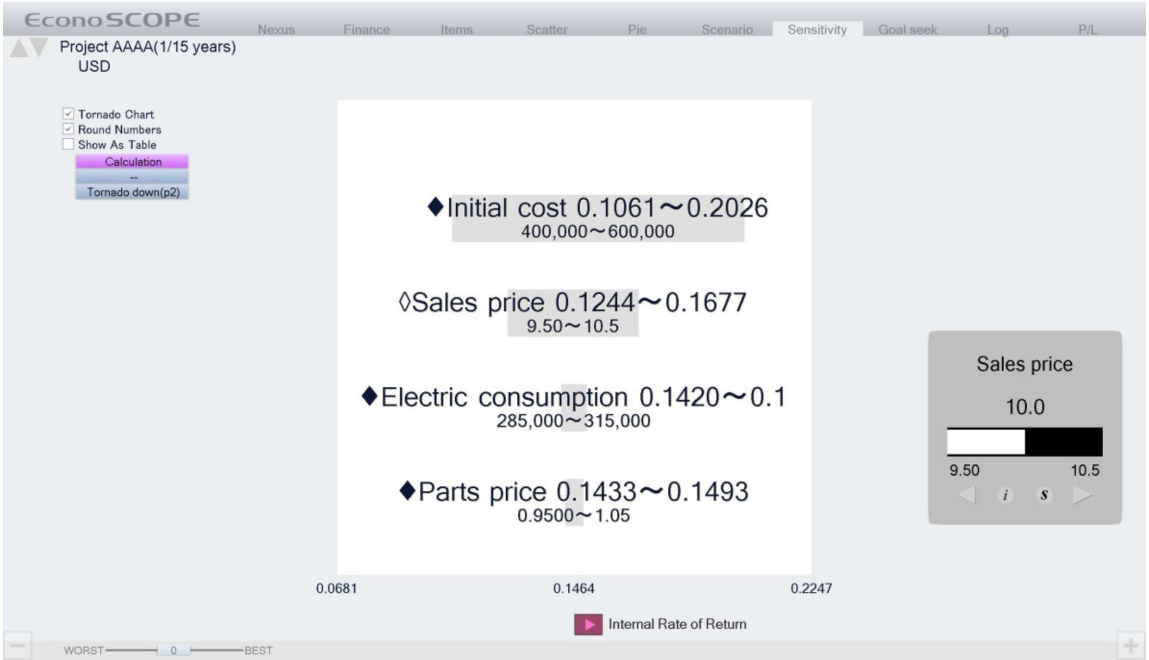


Figure 23 Screenshot of tornado chart

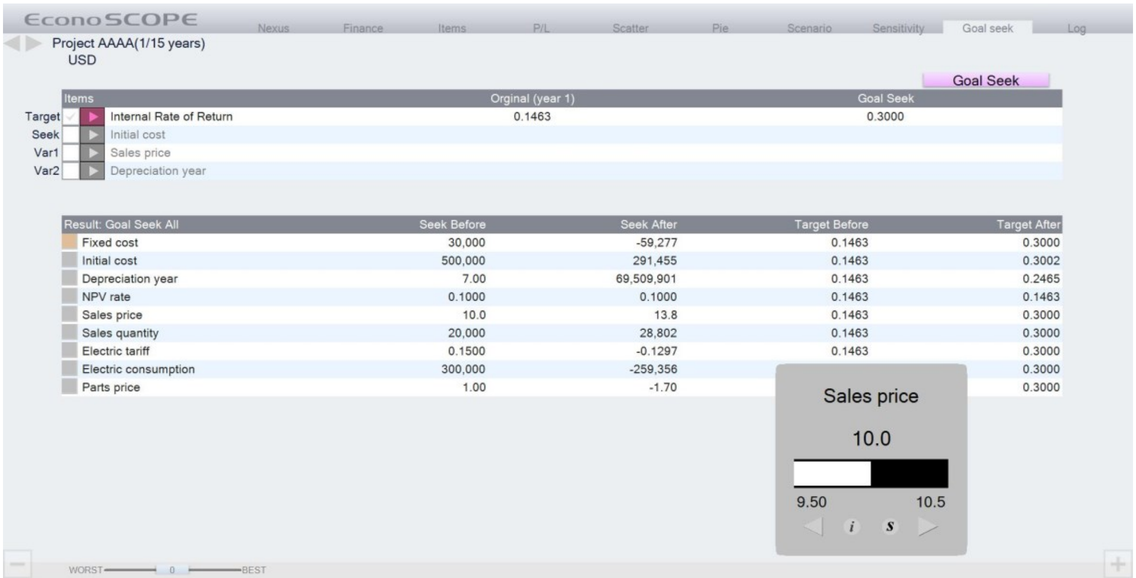


Figure 24 Screenshot of Goal Seek analysis

4.7 Validation Experiment

We performed a validation experiment with simple conditions to compare the method described in **Sections 4.4** to **4.7** with conventional spreadsheet software and assess the effectiveness of FDL.

4.7.1 Research goal and questions

In the experiment comparing FDL and conventional spreadsheet software, research goal (**RG**) and research questions (**RQs**) are defined as follows.

RG Investigate the advantages of FDL, compared with conventional spreadsheet software.

RQ4.1 Compared to conventional spreadsheet software, how efficiently can business people create financial calculation sheets with FDL?

RQ4.2 Compared to conventional spreadsheet software, how accurately can business people create financial calculation sheets with FDL?

RQ4.3 As the number of items whose value fluctuate increases, how efficiently business people can create financial calculation sheets with FDL and conventional spreadsheet?

RQ4.4 As the number of items whose value fluctuate increases, how accurately business people can create financial calculation sheets with FDL and conventional spreadsheet?

4.7.2 Subjects

The participants were 26 men and women ranging in age from 22 to 50 (average 29.3), who have experience in the use of spreadsheet software including use of the formula. Due to ethical issues, all the participants signed a form in which they accepted to participate in the experiment and let their data to be used in anonymized and aggregated form for research purposes. Their main work is planning, sales, marketing, and they do not know about FDL. Since the purpose of FDL is to widely modify the conventional spreadsheet work, we recruited participants using Excel formulas in business.

4.7.3 Variables and Hypotheses

Independent and dependent variables of this experiment are as follows:

Independent variables

Language / Tool:	FDL or conventional spreadsheet (Microsoft Excel 2010)
List of conditions:	Number of fluctuating items
Experiment order:	FDL → Excel or Excel → FDL
Participants' profile:	Programming experience, spreadsheet software experience, frequency to use spreadsheet software

Dependent variables

Time spent on creating calculation sheets
Time spent on creating graphs
Number of errors in calculation sheets

The hypotheses corresponding to **RQs** are as follows:

H4.1$_{null}$ The mean time spent for creating calculation sheet is longer or equal for the FDL than for the conventional spreadsheet software.

H4.1$_{alt}$ The mean time spent for creating calculation sheet is shorter for the FDL than for the conventional spreadsheet software.

H4.2$_{null}$ The mean number of errors in creating calculation sheet is greater or equal for the FDL than for the conventional spreadsheet software.

H4.2$_{alt}$ The mean number of errors in creating calculation sheet is less for the FDL than for the conventional spreadsheet software.

H4.3$_{null}$ As the number of items whose value fluctuate increases, the mean growth rate of the time spent to create a calculation sheet is greater or equal to that of FDL than in conventional spreadsheet software.

H4.3$_{alt}$ As the number of items whose value fluctuate increases, the mean growth rate of the time spent to create a calculation sheet is less to that of FDL than in conventional spreadsheet software.

H4.4$_{null}$ As the number of items whose value fluctuate increases, the mean of the differences in the number of errors when creating the calculation sheet is greater or equal for the FDL than for the conventional spreadsheet software.

H4.4$_{alt}$ As the number of items whose value fluctuate increases, the mean of the differences in the number of errors when creating the calculation sheet is less for the FDL than for the conventional spreadsheet software.

Table 11 shows the GQM (goal, question, metric) tree for these research goal, research questions, and hypotheses.

Table 11 GQM tree of this experiment

Goal	Question	Metric
RG	**RQ4.1** Efficiency	**H4.1** The time to create a calculation sheet is shorter for FDL than for Excel.
	RQ4.2 Accuracy	**H4.2** The number of errors in the created calculation sheet is smaller for FDL than for Excel.
	RQ4.3 Efficiency vs number of fluctuating items	**H4.3** The rate of change of the creation time when the number of items whose value fluctuates increases is smaller for FDL than for Excel.
	RQ4.4 Accuracy vs number of fluctuating items	**H4.4** The increment of the number of errors when the number of items whose value fluctuates increases is smaller in FDL than in Excel.

4.7.4 Experimental method

The FDL description methods needed to perform the tasks assigned were first described for about 30 minutes and the experiment was then conducted for three types of conditions. In the experiment, the participants were presented with a list of calculation conditions like the one shown in Listing 6.1 and were asked to create FDL and a spreadsheet using Microsoft Excel 2010 for the same content. Then, using both developed tool and Excel's functions, the participants were asked to create line graphs, bar graphs, pie charts, and bubble charts and to perform a Goal Seek. Spider charts and tornado charts were excluded from measurement since preliminary experiments revealed cases where creation of such a chart took a long time or was somewhat difficult. In order to eliminate learning effects, participants were divided into two groups. Participants in the first group are required to create in the order of Excel sheet and then FDL in tasks 1 and 3, while in task 2, in order of FDL and then Excel sheet as shown in **Table 12**. The participants in the second group were asked to create in reverse order. The time taken for (1) creation of FDL sheet, (2) creation of FDL sheet and graph (developed tool), (3) creation of Excel calculation sheet, and (4) creation of Excel calculation sheet and graph, were measured and compared. As shown in the underlined part of the **Listing 1**, tasks 2 and 3 have one and two items whose value fluctuate for specifying the minimum and maximum values, respectively. Therefore, participants are asked to combine them and calculate 2 times and 4 times, respectively.

Listing 1 Example of list of conditions (task-3)

Period: 20 years
Sales: No. of software packages sold × unit selling price
Maintenance cost: No. of maintenance personnel × unit personnel cost
Merchandizing cost: advertising cost + catalog production cost
Depreciation: depreciation of initial development cost over 5 years (straight-line method)
Total expenditure: maintenance cost + merchandizing cost + depreciation
Profit: Sales minus total expenditure
Total profit: Total of annual profit
IRR: Calculated as recovery of initial development cost through profits
NPV: Calculated as recovery of initial development cost through profits. rate = 1.5%
No. of software packages sold: 10,000 first FY progressing at annual rate of 110%
Unit selling price: $50 first FY progressing at annual rate of 95%
Unit personnel cost: $50,000 first FY progressing at annual rate of 105%
No. of maintenance personnel: 8
Initial development cost: minimum $100,000, maximum $200,000
Advertising cost: minimum $20,000, maximum $40,000
Catalog production cost: $2,000

Table 12 Three tasks of validation experiment

	task1	task2	task3
group 1 (13 participants)	Excel->FDL	FDL->Excel	Excel->FDL
group 2 (13 participants)	FDL->Excel	Excel->FDL	FDL->Excel
number of fluctuating items	0	1	2

4.7.5 Experimental results

The results of measuring the time required for creating FDL sheet, a conventional spreadsheet, and graphs are shown in **Figure 25**. As shown, the time required to create FDL sheet was about 60% that of creating a Excel's spreadsheet on average. In addition, considering the time required to create graphs using spreadsheet software, it can be seen that FDL and its visualization and analysis tools are more effective. In fact, after creating the FDL sheet, the time to display each graph with the developed tool is several seconds of item selection only. In the following figures, the upper and lower ends of the box are 75% and 25% of the data distribution, and the middle line shows the median value. The black circle indicates the average value, and the upper end and the lower end of the whisker indicate the maximum and minimum values. "(*)" indicates that the p value in the T test is less than 0.01. The tables show the mean, the standard deviation (SD), the median and the p-value of the T test.

As shown in **Table 13**, the averages of the creation time of all cases are shorter for FDL than for Excel. The p values of T test were less than 0.01 for all tasks including graph creation except for Task 1 in the calculation sheet creation. This result allows us to reject the hypothesis **H4.1$_{\text{null}}$**, and accept the alternative hypothesis **H4.1$_{\text{alt}}$**.

Next, the number of errors that occurred for a total of 78 tasks (26 participants times 3 tasks) was found to be 221 and 37 for the Excel and FDL, respectively. **Figure 26** and **Table 14** show the number of errors per participant. Errors seen in FDL sheets were mistakes in numerical inputs and operator inputs. There was no mistake in the entry of reference which can be seen much in Excel. The average number of errors is 0.5 to 4.5 more in Excel than FDL. The p value of the T test was less than 0.01 on tasks 2, 3 and on average. Therefore, we can reject the hypothesis **H4.2$_{\text{null}}$**, and accept the alternative hypothesis **H4.2$_{\text{alt}}$**.

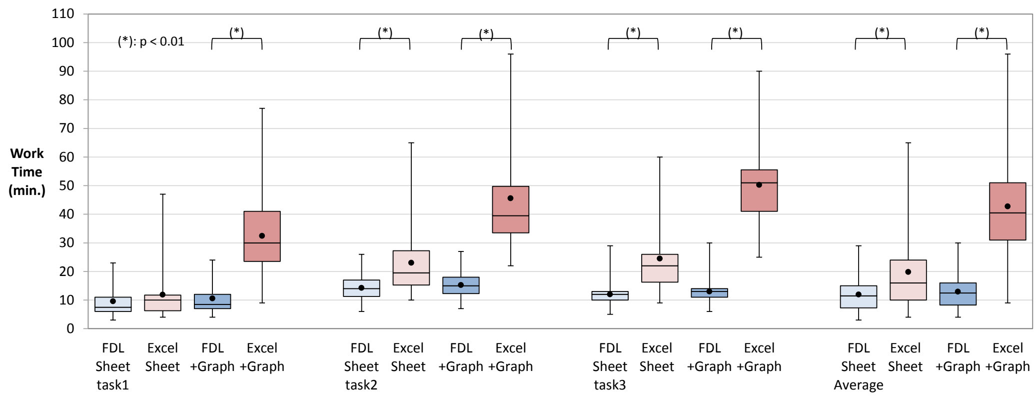


Figure 25 Work time measured for creating FDL and Excel's spreadsheet, and both sheets and graphs creation.

Table 13 Mean, standard deviation, median, and t-Test p for the work time measured for creating FDL and Excel's spreadsheet, and both sheets and graphs creation.

	task1				task2				task3				Average			
	Sheet		Sheet+Graph		Sheet		Sheet+Graph		Sheet		Sheet+Graph		Sheet		Sheet+Graph	
	FDL	Excel	FDL	Excel	FDL	Excel	FDL	Excel	FDL	Excel	FDL	Excel	FDL	Excel	FDL	Excel
Mean	9.5	11.9	10.5	32.4	14.2	23.0	15.2	45.6	12.0	24.5	13.0	50.2	11.9	19.8	12.9	42.7
SD	5.3	10.8	5.3	15.1	5.1	12.7	5.1	18.8	4.9	12.6	4.9	14.2	5.4	13.2	5.4	17.7
Median	7.5	10.0	8.5	30.0	14.0	19.5	15.0	39.5	12.0	22.0	13.0	51.0	11.5	16.0	12.5	40.5
t-Test p	2.1E-01		8.2E-09(*)		1.0E-03(*)		1.7E-09(*)		4.2E-07(*)		4.3E-15(*)		1.8E-08(*)		2.5E-28(*)	

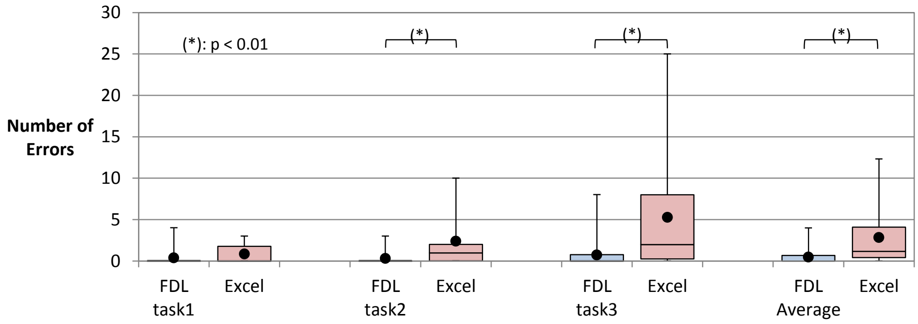


Figure 26 Number of errors per participant.

Table 14 Mean, standard deviation, median, and t-Test p for the number of errors per participant.

	task1		task2		task3		Average	
	FDL	Excel	FDL	Excel	FDL	Excel	FDL	Excel
Mean	0.38	0.85	0.31	2.38	0.73	5.27	0.47	2.83
SD	1.13	1.08	0.74	3.06	1.73	6.46	0.92	3.31
Median	0.00	0.00	0.00	1.00	0.00	2.00	0.00	1.17
t-Test p	1.1E-02		1.2E-03(*)		1.3E-03(*)		3.8E-04(*)	

The learning effect is one of the threats of internal validities. For each task, the same content is created in both FDL and Excel, so the order may affect the results. In order to alleviate this effect, participants were divided into two groups, and experiments were conducted in reverse order. Also, a comparison between the time of the task executed in the order of Excel → FDL (i.e., tasks 1 and 3 of the first group and task 2 of the second group) and the time in the order of FDL → Excel is shown in the **Figure 27**. As shown in **Table 15**, here was no significant difference in each order.

On the other hand, the difference in the number of errors was significant in the case of Excel as shown in **Figure 28** and **Table 16**. However, this is the opposite effect from what is assumed by the learning effect. Since tasks using Excel take longer than using FDL, it may be because the concentration of participants has dropped. In this experiment, these effects can not be separated and new control experiments are necessary in the future.

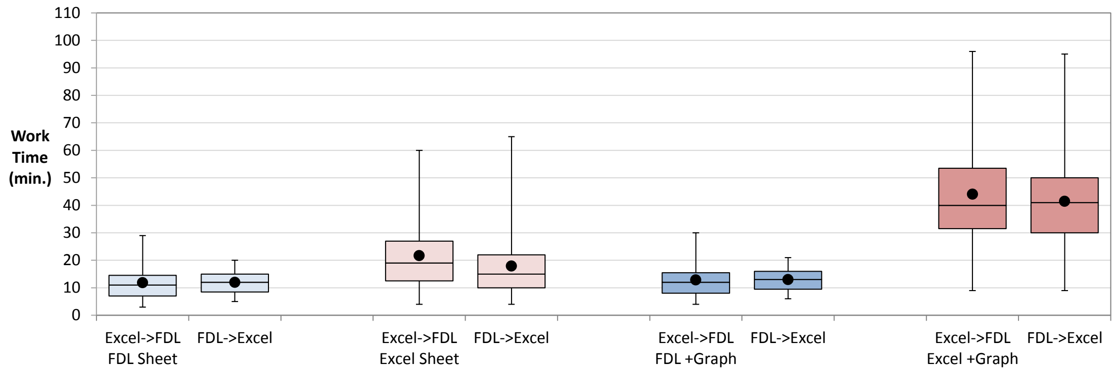


Figure 27 Comparison of working time by order of experiment.

Table 15 Mean, standard deviation, median, and t-Test p for working time by order of experiment.

	FDL Sheet		Excel Sheet		FDL Sheet + Graph		Excel Sheet + Graph	
	Excel->FDL	FDL->Excel	Excel->FDL	FDL->Excel	Excel->FDL	FDL->Excel	Excel->FDL	FDL->Excel
Mean	11.8	12.0	21.7	17.9	12.8	13.0	44.1	41.4
SD	6.4	4.3	12.9	13.4	6.4	4.3	19.3	16.0
Median	11.0	12.0	19.0	15.0	12.0	13.0	40.0	41.0
t-Test p	9.2E-01		1.5E-01		9.2E-01		3.7E-01	

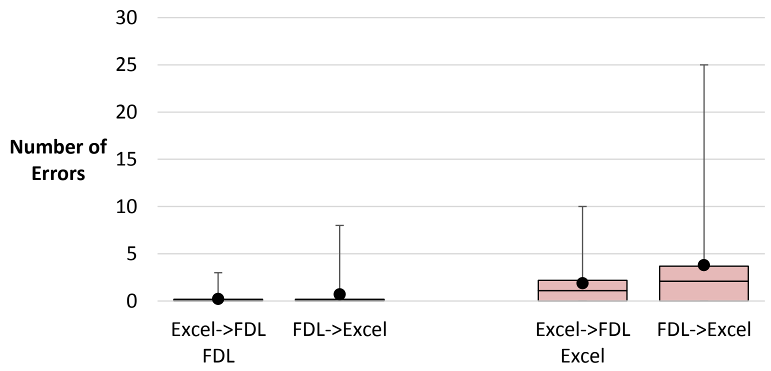


Figure 28 Comparison of error number by order of experiment.

Table 16 Mean, standard deviation, median, and t-Test p for error number by order of experiment.

	FDL		Excel	
	Excel->FDL	FDL->Excel	Excel->FDL	FDL->Excel
Mean	0.23	0.72	1.87	3.79
SD	0.63	1.65	2.90	5.56
Median	0.00	0.00	1.00	2.00
t-Test p	5.8E-02		4.9E-3(*)	

The main differences between tasks 1, 2, and 3 are the number of items whose values fluctuate, and their numbers are 0, 1, and 2, respectively. **Table 17** shows how much the work time has changed when the number of items whose values fluctuate from 0 to 1(from task1 to task2) and 1 to 2 (from task2 to task3). As shown in **Table 17**, except for calculation sheet and graph creation of task 1 to task 2, work time growth rate of Excel is larger than FDL. However, there is only one case where p value of T test is less than 0.01. Therefore, we can not conclude that we can reject hypothesis **H4.3$_{\mathbf{null}}$**.

Table 18 shows the increment of the number of errors when the number of items whose values fluctuate increases. In both cases the increment of the number of Excel errors is greater than FDL, but in one case the p value is not less than 0.01. So the hypothesis **H4.4$_{\mathbf{null}}$** can not be rejected.

Table 17 Work time growth rate

	task1 -> task2				task2 -> task3			
	Sheet		Sheet+Graph		Sheet		Sheet+Graph	
	FDL	Excel	FDL	Excel	FDL	Excel	FDL	Excel
Mean	178%	269%	167%	157%	90%	115%	90%	116%
SD	86%	176%	74%	60%	36%	43%	33%	27%
Median	176%	231%	167%	140%	83%	102%	84%	116%
t-Test p	3.3E-02		5.6E-01		4.5E-02		5.9E-3(*)	

Table 18 Increment of the number of errors

	task1 -> task2		task2 -> task3	
	FDL	Excel	FDL	Excel
Mean	-0.08	1.54	0.42	2.88
SD	0.89	2.35	1.55	4.36
Median	0.00	0.50	0.00	1.50
t-Test p	8.8E-3(*)		1.1E-02	

4.7.6 Subjective evaluation

After completing the task, the participants were asked to fill in the questionnaires as shown in **Listing 2**. The assessment relies on a four-point Likert-Scale ranging from “Completely agree” (1), “Agree” (2), “Disagree” (3), and “Completely disagree” (4).

Listing 2 Questionnaires

Compared to conventional spreadsheet software,
(a) FDL can save time.
(b) FDL reduces mistakes such as erroneous input.
(c) FDL makes it easier to understand sheets made by others.
(d) FDL makes it easier for others to understand sheets created by you.
(e) FDL has high reusability.
(f) FDL is easy to use.

Questionnaires **(a)** and **(b)** are subjective evaluations corresponding to **RQ4.1** and **RQ4.2**. As shown in **Figure 29**, both show high scores. That is, from both objective measurement data and subjective judgment, it can be said that FDL can reduce time and error compared to conventional spreadsheet. On the other hand, questionnaires **(c)** to **(f)** correspond to the following other **RQs**. These also show relatively high scores, and will be investigated together with objective measurements in future research.

RQ4.5: Compared to conventional spreadsheet software, how easy can business person understand financial calculation sheets with FDL created by others?
RQ4.6: Compared to conventional spreadsheet software, how easy can business person make others understand the finance calculation sheets with FDL he/she created?
RQ4.7: Compared to conventional spreadsheet software, how much reusability are the financial calculation sheets that business person created with FDL?
RQ4.8: Compared to conventional spreadsheet software, how easy to use FDL and visualization / analysis tool?

In addition, participants were asked to write comments with free description. They commented that it is difficult to memorize operators of FDL, and it is preferable to use a menu. On the other hand, there were positive comments such as "can be described intuitively", "input is simple", and “check of mistakes is easy".

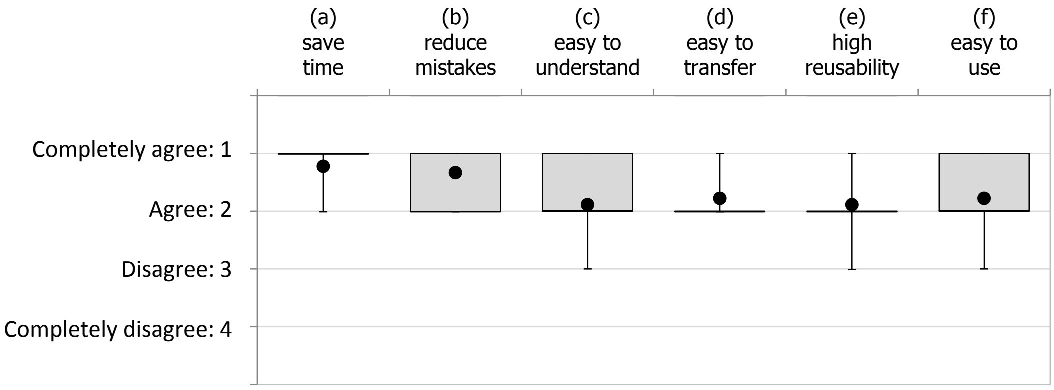


Figure 29 Result of the questionnaire

4.7.7 Threats to validity

In the experiments, there may be several threats to validity. The internal validity needs to be verified whether the control parameters are correctly reflected in the hypotheses judgment. The external validity must be verified whether the hypotheses can be generalized.

Task size can be another threat to internal validity. Each task of creating sheets and graphs in both FDL and Excel is designed to finish in about an hour. Actually, the average time required for each task was about 55 minutes. Also, after completing each task, the participants took a break for more than 10 minutes.

The task order may also be threat to internal validity. Experiments were conducted in the order of task 1, task 2, and task 3. The number of items whose values fluctuate for these tasks is 0, 1, and 2. The task order may affect **RQ4.3** and **RQ4.4**, but in the preliminary experiments, there was no difference depending on the order.

Subject's experience is a typical example that can be threat to external validity. Seventeen of the 26 participants had experience of creating Excel macros. As a result of comparison with the experience of macro creation, no significant difference was found in time and number of errors as shown in **Table 19** and **Table 20**. Likewise, there was no correlation between years of Excel's experience, frequency of use, age, etc. and the results.

Table 19 Comparison of working time with macro experience

	FDL Sheet		Excel Sheet		FDL Sheet + Graph		Excel Sheet + Graph	
Excel's Macro experience	yes	no	yes	no	yes	no	yes	no
Mean	11.6	12.5	19.0	21.2	12.6	13.5	40.9	46.2
SD	5.3	5.7	14.0	11.8	5.3	5.7	18.2	16.3
Median	11.0	12.0	16.0	19.0	12.0	13.0	38.0	44.0
t-Test p	5.0E-01		4.9E-01		5.0E-01		2.1E-01	

Table 20 Comparison of number of errors with macro experience

	FDL		Excel	
Excel's Macro experience	yes	no	yes	no
Mean	0.47	0.48	2.51	3.44
SD	1.32	1.19	4.00	5.38
Median	0.00	0.00	1.00	2.00
t-Test p	9.7E-01		3.9E-01	

4.7.8 Where FDL is most effective

Based on our results, FDL can create a finance calculation sheet and draw graphs more efficiently and accurately than conventional spreadsheet software. It tends to be effective especially when there are many items whose values fluctuate. We also believe that it is effective for a wide user regardless of technical experience. Furthermore, since it seems to be effective when sharing calculation sheets with colleagues and customers as shown in **RQ4.5** to **RQ4.7** and **Figure 29**, objective measurement data on information transmission will be gathered and verified in the future like Kieburts' work [119].

4.8 For consensus formation in business

Spreadsheets have taken on an indispensable existence in all sorts of business activities, but as their content becomes increasingly complex and larger in scale, the difficulty of understanding that content and the occurrence of errors have become issues. Our aim in developing FDL that clearly separates input and output and operators and operands was to provide a description system that could enhance legibility and make it easier to understand spreadsheet content. The FDL format can be converted to the conventional spreadsheet format and general purpose programming languages such as C++. At the same time, there is no advantage to describing in FDL a spreadsheet consisting only of numerical values with almost no formulas.

This chapter described our original visualization and analysis tools, but ordinary spreadsheet software functions may be described as-is as operators separately from operands (item names) in FDL format and a program for calculating those functions may be developed.

We developed FDL for use by the many users familiar with spreadsheet software formulas without the need for programming skills using high-level languages like C++ and Java. At present, more than 300 people are using FDL and its visualization and analysis tools. These users have commented that in addition to shortening development time compared with conventional spreadsheet software, FDL facilitates analysis that in the past they had given up on. Finally, experiments have shown that FDL is more effective and efficient than conventional spreadsheet software in terms of work time and error occurrence rate.

As described in **Section 4.3** , current FDL has DAG constraints. However, in complex business situations, causal relationships between items can be cycled. Therefore, we plan to extend FDL to make it possible to describe such a circular calculation.

Chapter 5 Decision making process of individuals: Optimal solution and exploration

Summary of Chapter 5

In this chapter, I describe the modeling of individual's decision making process including irrational judgment when going through the maze of virtual space. The subject is instructed to advance as straight forward as possible toward the front of the start position. A total of 8 tasks were performed on the four types of 3D models when only the subjects operate or the software intervenes to change the uncertainty. The distance between the final arrival point and the correct goal, the judgment time, and the ratio of unreasonable choice away from the correct direction did not differ from the difference in uncertainty. However, when modeling with Bayesian inference and *inverse Bayesian inference*, the higher the uncertainty, the more the *inverse Bayesian inference* tends to be able to estimate with higher accuracy than Bayesian inference. Immediately after the start of the experiment, the subject was in the stage of the shrinking process, making decisions that seemed to be the optimal solutions. Subsequently, the subject suddenly noticed that it departed from the goal, and switched to an expanding process to explore for a way to return. In the case of high uncertainty, since such cases frequently occurred, it is considered that the accuracy of the *inverse Bayesian inference* was high. Thus, when the uncertainty is high and the decision-making method itself changes suddenly, it is shown that modeling by *inverse Bayesian inference* is effective. This chapter is based on [120].

5.1 Sense of direction in virtual 3D space

When moving or changing direction, humans continue to update their own local coordinate system to the world coordinate system. Liberman rotated the goldfish in the direction of the two axes orthogonal to each other, and then observed the nerve cells of the goldfish which became impossible to swim. Before rotation, the microfilaments of the cytoskeleton parallel to the dendrites collapsed after rotation, indicating that decision-making to select the coordinate system is also done at the cellular level [121] [122].

We designed this experimental system to make visual information presented to the subject and the results of the subject's button operations both incomplete and uncertain so that decision-making might change suddenly. A total of 418 trials were performed in the experiment, and it was often found that decision-making would change abruptly. We model these decision processes using Bayesian inference and inverse Bayesian inference and compare them.

In this experiment, the information presented to the subjects is uncertain, and the pattern of decision of the subjects is designed to change suddenly. The uncertainty of the information obtained in the actual decision making depends on the situation at that time and the knowledge of the subject. In addition, even if it is possible to make a reasonable judgment with sufficient time and necessary knowledge, in actual circumstances, there are many cases where judgment must be made immediately. Therefore, we designed an experiment that (i) does not rely on past knowledge and experience, (ii) reproduces the situation to be judged one after another, and (iii) the degree of uncertainty can be controlled.

5.2 Materials and methods

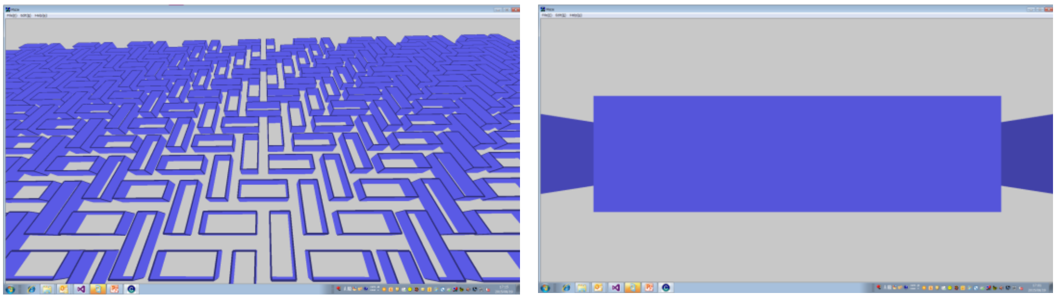
5.2.1 Design

In this experiment, the decision-making process for moving to a certain destination was reproduced in a virtual space. Even in the real world, by misunderstanding the target direction by repeating turning left or right several times, or misunderstanding the estimation of the distance traveled, we may lose a sense of direction. In order to reproduce such a situation, we developed software that proceeded in the virtual space by right and left selection of subjects. In this experiment, (i) subjects do not need any special prior knowledge, and (ii) they need to continue deciding which way to proceed. Also, (iii) the degree of uncertainty is controlled by the forced selection of the software. The left and right data series selected by the subject are analyzed using Bayesian and BIB inference shown in **Section 5.2** with binary values of 1 and -1, respectively.

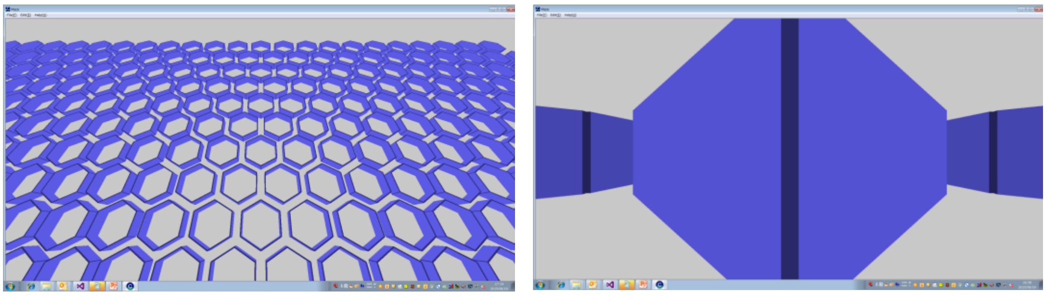
5.2.2 Procedure and tasks

As shown in the **Figure 30 (left)** and **Table 21**, there are four types of three-dimensional (3D) objects: (a) *rectangles*, (b) *hexagons*, (c) *walls created in ad hoc*, and (d) *pillars*. These objects are arranged within a prescribed area. The 3D objects (a) to (c) were assumed to test the sense of direction, and the 3D object (d) was assumed to test the sense of migration. As shown in **Figure 30 (right)**, the space viewed from the first person's perspective consisting of 3D objects is displayed on a computer screen. The viewpoint position in 3D space moves according to the subject's left/right key selection. In the case of (a) to (c), the view direction is also changed. In the case of (c), walls are automatically created on the left and right in the next dead end (impossible in reality). In the case of (d), the pillars-like objects are randomly arranged, and the viewpoint advances in a zigzag manner diagonally forward by selection of the left and right of the subject. The migration length in the left or right (lateral) direction is set randomly within a certain range while the migration length in the forward direction is fixed. In all cases, the time to move to the point where the next selection is required is 1000 milliseconds. In tasks 1 to 4, only the subject selects the direction of movement. To increase the degree of uncertainty, in tasks 5 to 8, the software automatically selects the left or right key once for each of the three key selections.

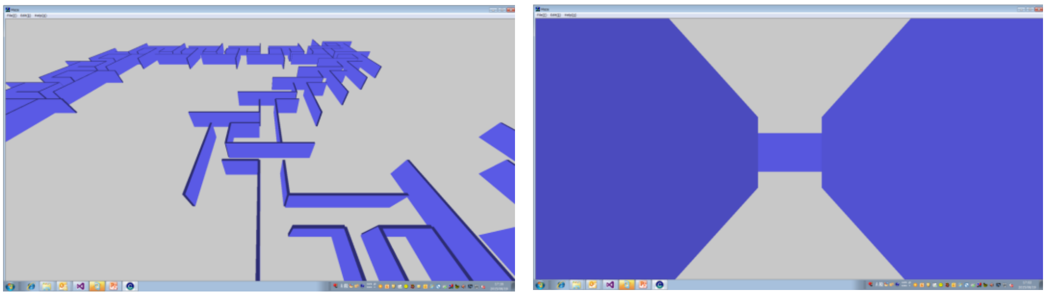
If the same experiment were to be performed in the third person's perspective such as in **Figure 30 (left)**, the subject would likely arrive at the correct goal position with ease. This is because such a view would enable the subject to determine the relationship between the start and goal positions and his or her current location at all times. However, in the first persons' perspective of this experiment as shown in **Figure 30 (right)**, the subject cannot look backwards and can only obtain information in the forward direction within a certain field of view angle (40 degree in this experiment). So, the information presented to the subject is uncertain in this experiment.



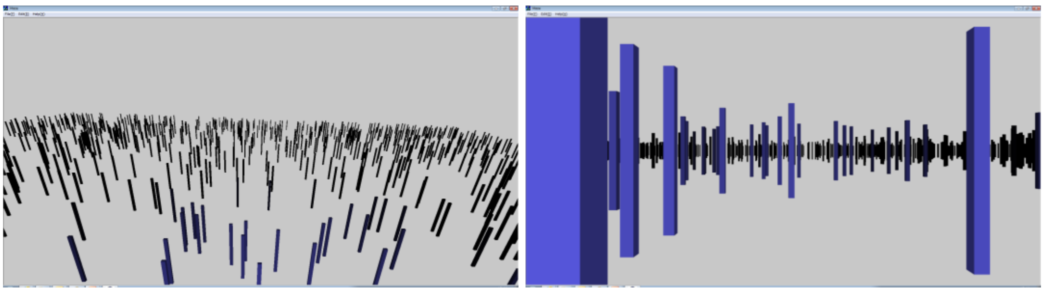
(a) Rectangle



(b) Hexagon



(c) Ad hoc walls



(d) Pillar

Figure 30 Screenshot of experiment on moving through a virtual three-dimensional space arranged with objects. The viewpoint in this three-dimensional space moves according to the subject's selection of a left or right key. The subjects were instructed to move in the forward direction from the start position. Key-operations and the elapsed time were recorded.

Table 21 Task list of this experiment. For tasks 1 to 4, only the subject selects direction of movement, while for tasks 5 to 8, the software randomly selects left and right, once in three times. The subjects perform all tasks 1 – 3 times.

Task	3D model	selection by
1	(a) Rectangles	subject only
2	(b) Hexagons	
3	(c) Ad hoc walls	
4	(d) Pillars	
5	(a) Rectangles	subject->subject->software
6	(b) Hexagons	
7	(c) Ad hoc walls	
8	(d) Pillars	

5.2.3 Variables and research questions

Independent and dependent variables of this experiment are as follows:

Independent variables

- 3D models (*rectangles*, *hexagons*, *ad hoc walls*, and *pillars*)
- Selection left or right by subjects and software (subject only and both subject and software)

Dependent variables

- Left or right selection sequence
- Time required for judgment
- Number of unreasonable judgements (straying from the centerline connecting the start and goal positions)
- Goal position

The research questions of this experiment are as follows:

RQ5.1 Does the subject's goal position change depending on the 3D model?
RQ5.2 Does time of subject's judgment change depending on 3D model?
RQ5.3 Does the ratio of subject's unreasonable judgment change depending on 3D model?
RQ5.4 Do the subject's goal position, the ratio of irrational judgment, and the judgment time change with the degree of uncertainty?

RQ5.5 Can Bayesian inference and BIB inference model actual data?

RQ5.6 Does the accuracy of Bayesian inference and BIB inference change with the degree of uncertainty?
RQ5.7 Does BIB inference provide higher accuracy than Bayesian inference?
RQ5.8 Does the difference in error between Bayesian inference and BIB inference change as the degree of uncertainty increases?

5.2.4 Participants and instruction

The subjects were naïve college students numbering 30 in total (18 men and 12 women) without any detailed knowledge of this field. The subjects were instructed that “the goal lies straight ahead—select either the left or right arrow key to advance toward the goal” and were encouraged to move straight ahead as much as possible. In each trial, the subjects performed these key operations until 3D objects could no longer be seen, which took about 50 steps. In tasks 3 and 7 (*ad hoc walls*), each experiment was continued up to 50 steps. Each task was completed in about 1 minute and a break of about 30 seconds was set up until the next task execution. The order of the tasks was 4, 8, 1, 5, 2, 6, 3, and 7, and the subjects were conducted 1 to 3 times. There was no notable difference between the same tasks of the same subjects, and no influence of fatigue and accustomed was observed. Due to ethical issues, all the participants signed a form in which they accepted to participate in the experiment and let their data to be used in anonymized and aggregated form for research purposes.

5.3 Results

5.3.1 Result of experiments

Examples of experimental results are shown in **Figure 31**. The line graph in the figure represents the locus of viewpoint motion when observing the 3D space from above. The vertical direction in this graph corresponds to forward direction in the experiment. The points on this locus indicate locations where left or right was selected, and the numerical values indicate elapsed time (second) since the start of the experiment.

The distribution of the distance between the goal position and the actual arrival point for each task is shown in **Figure 32**. The results of the experiment using *pillars* (tasks 4 and 8) are significantly smaller than that of the other 3D objects (**RQ5.1**). **Figure 33** shows the distribution of time until the subject selects left and right keys. In the tasks of the *pillars*, the time required for the decision is significantly shorter (**RQ5.2**).

The result in straying from the centerline connecting the start and goal positions can be seen.

Figure 34 shows the percentage of such an unreasonable judgment among left and right selection by subjects in about 50 steps in each trial. This percentage does not include selection by software. The proportion of unreasonable judgment when the 3D object is *rectangles* and *hexagons* is higher than that of *ad hoc* and *pillars* (**RQ5.3**).

As shown in **Figure 32** to **Figure 34**, these three results (goal position, judgment time, and percentage of unreasonable judgement) were unrelated to the presence or absence of forced selection by the program except task 4 & 8 in unreasonable decision (**RQ5.4**). In tasks of *pillars*, the subject reaches a position not far from the correct goal, the unreasonable judgment rate is low, and the judgment time is short. Therefore, it can be said that these are the easiest tasks for the subject. That is, the influence of the *change of the distance* on the task is smaller than the *change of the direction*. In tasks of *rectangles* and *hexagons*, the distance from the correct goal was large, the ratio of unreasonable judgment was high, and it took time to judge, so these may be difficult tasks for the subject. A possible reason is that in the case of *rectangles,* there are two types of distances traveled by left or right selection. Also, in the case of *hexagons*, turning 60 degrees is rarely experienced in reality. For these reasons, it may be difficult in these two cases.

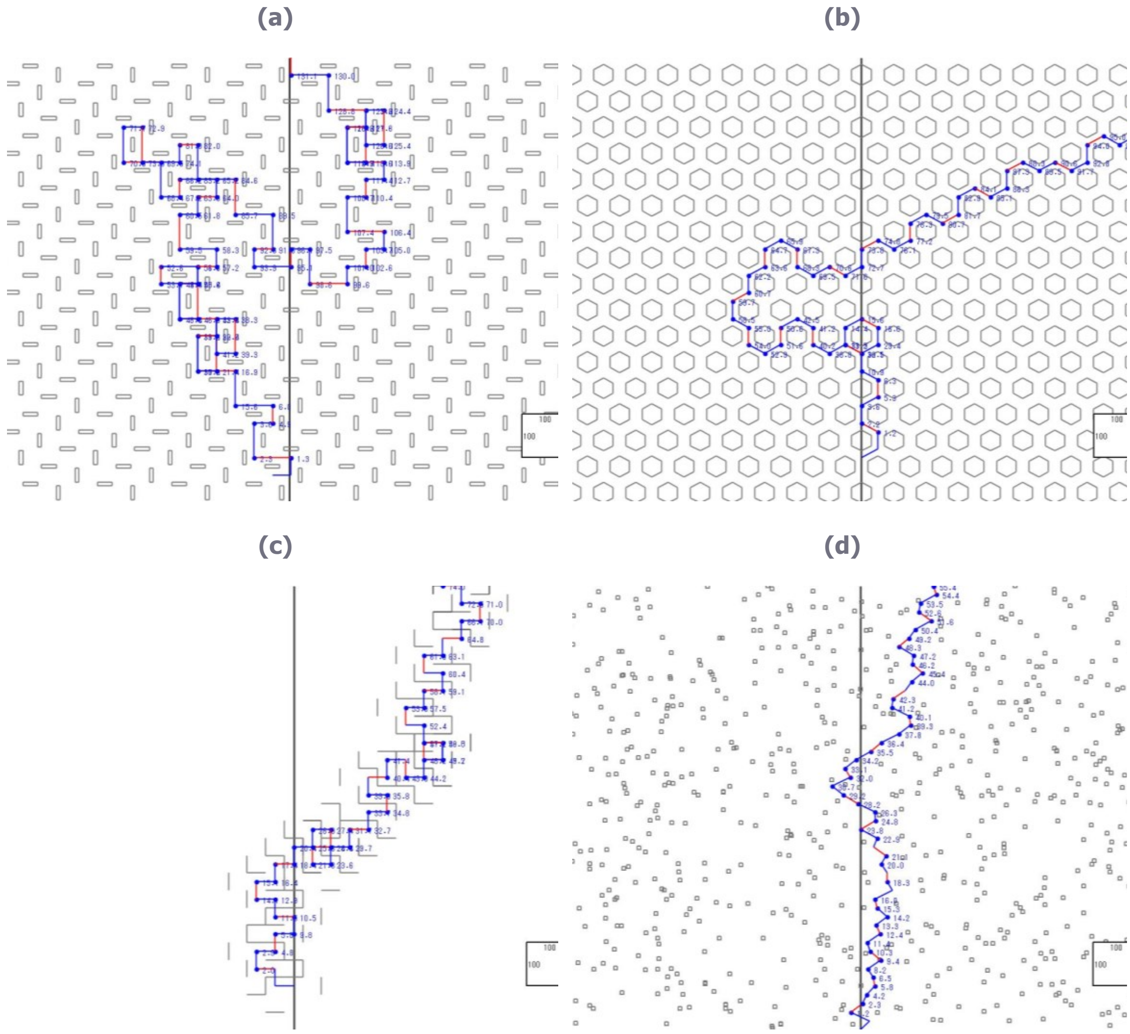


Figure 31 Bird's-eye view of experimental results showing movement history. (a) task 5, (b) task 6, (c) task 7, and (d) task 8. The line graph is the locus of viewpoint motion when observing the three-dimensional space from above and the vertical direction corresponds to forward direction. Points on the locus indicate locations where left or right was selected, and the numerical values indicate elapsed time (second) since the start of the experiment. The red line shows the trajectory selected by the software. The lower right square shows the scale of the space, and the length in the longitudinal direction (goal direction) and the length in the lateral direction are both 100.

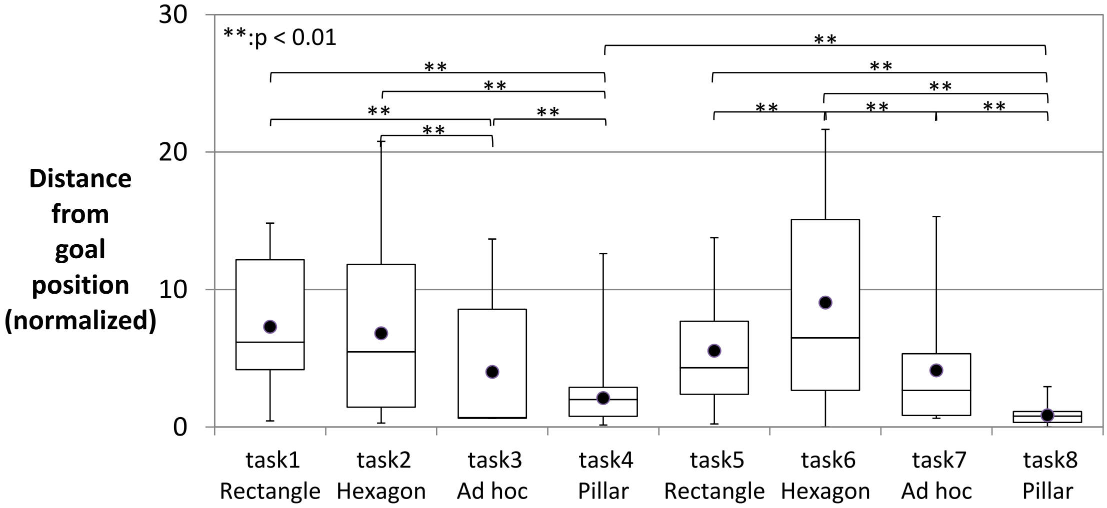


Figure 32 Distributions of displacement from the correct goal position at the end of the test. The value of the vertical axis is normalized by the average value of the moving distance of each step. Variances of tasks 4 and 8 ((d) *pillars*) are significantly smaller. In this figure (also in Figure 33, Figure 39, and Figure 42), the upper and lower ends of the box are 75% and 25% of the data distribution, and the middle line shows the median value. The black circle indicates the average value, and the upper end and the lower end of the whisker indicate the maximum and minimum values. The number of trials for tasks 1 to 8 are 49, 51, 50, 58, 51, 52, 50, and 57, (total 418 trials) respectively.

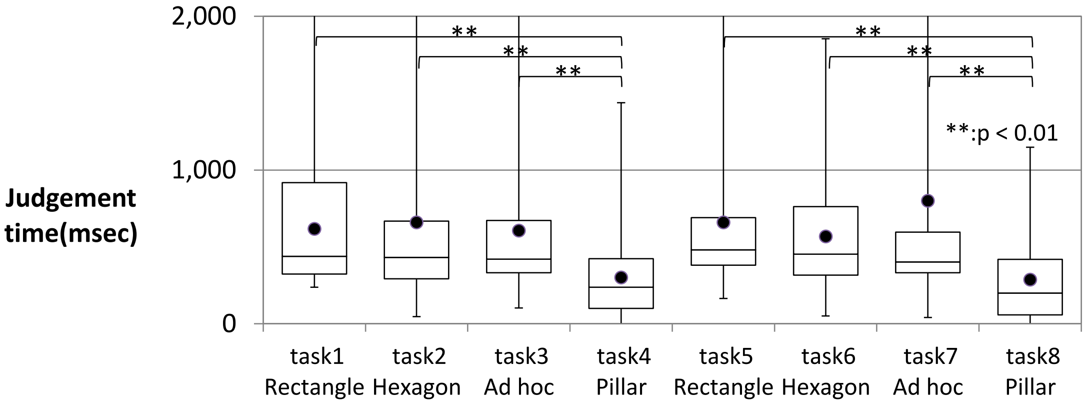


Figure 33 Time taken to select left and right of the subject (millisecond). Time for tasks 4 & 8 ((d) *pillars*) was significantly shorter.

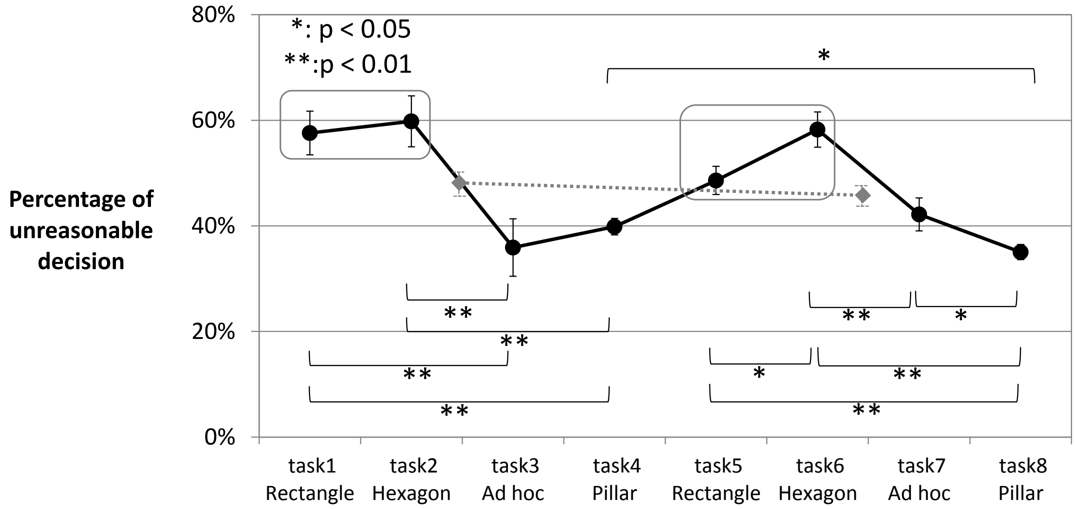


Figure 34 Percentage of unreasonable judgment among subjects' decisions in each test. Error bars indicate standard error. Tasks 1, 2, 5, and 6 ((a) *rectangles* & (b) *hexagons*) are higher than tasks 3, 4, 7, and 8 ((c) *ad hoc* & (d) *pillars*) in terms of unreasonable judgment. The dotted line shows the average value of task 1-4 and task 5-8.

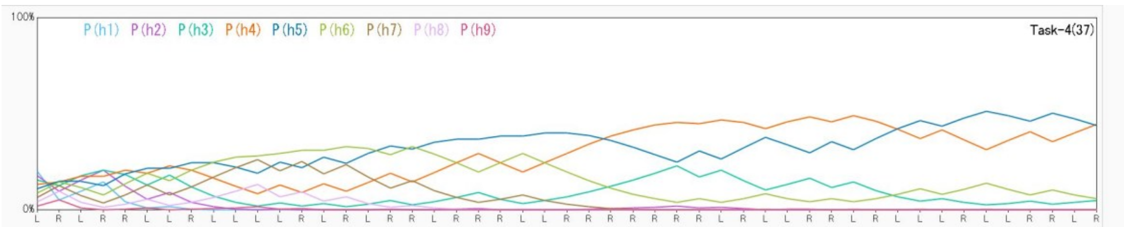
5.3.2 Evaluation of results using Bayesian inference and BIB inference

In this section, we calculate Bayesian inference and BIB inference described in **Section 5.2** setting $d^t = 1$ when left is selected and $d^t = -1$ when right is selected in the experiment. The number of hypotheses N is set to 9. **Figure 35 (a)** shows the prior probability $P^t(h_i)$ of hypothesis h_i against time obtained only by equations (1) and (2) (i.e. only Bayesian inference), and **Figure 35 (b)** shows those obtained by equations (1) to (4) (i.e. BIB inference). The probability where $d^t = 1$, that is, likelihood $P^t(d|h_i)$ of each hypothesis by Bayesian inference is invariable but likelihood by BIB inference is updated sequentially as shown in **Figure 36**.

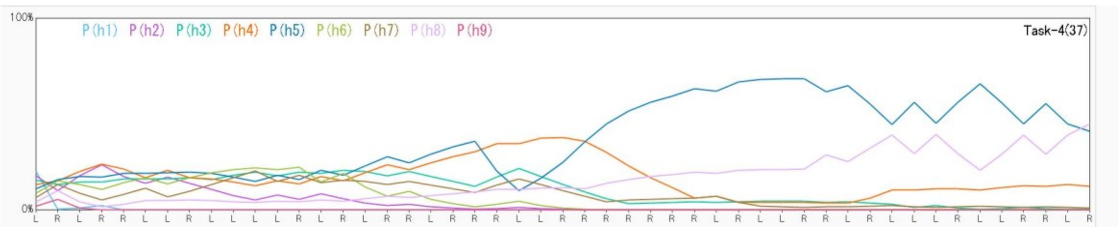
Figure 37(a) shows average from start time $P^t_{ac}(d)$ for percentage occurrence of d^t=1 and the likelihood of Bayesian inference $P^t(d|h_i)$ for the hypothesis having maximum prior probability $P^t(h_i)$, and **Figure 37(b)** shows moving average $P^t_{mov}(d)$ and likelihood of BIB inference. The results of Bayesian inference follow average $P^t_{ac}(d)$ from start time while the results of BIB inference follow moving average $P^t_{mov}(d)$. It can therefore be said that Bayesian inference is effective for a steady data-appearance pattern and that BIB inference is effective for a suddenly changing data-appearance pattern.

Next, in order to further investigate whether actual behavior can be modeled, comparison was made based on the cumulative value of the likelihood with the highest prior probability. Here, we compared cumulative change of experimental data $c_{raw}{}^t = \sum_{k=0..t} d^k$ with cumulative change of the results obtained by Bayesian inference and BIB inference expressed as $c_{inf}{}^t = \sum_{k=0..t} (2P^k(d|h_i) - 1)$. Here, since $P^t(d|h_i)$ takes a value from 0 to 1, it is converted from -1 to 1. As shown by the example in **Figure 38**, Bayesian inference cannot track the sudden change in behavior occurring near the middle of the experiment while BIB inference can.

Table 22 and **Figure 39** show the mean square error $E = \sum_{t=0..T-1}(c_{raw}{}^{t} - c_{inf}{}^{t})^2/T$ between the actual data and the inference results. $E^{1/2}$ in **Table 22** means *times* where the inference result is different from the actual selection. That is, Bayesian inference can be estimated within the error range of 0.80 to 5.52 times, and in the BIB inference it can be estimated within the error range of 0.80 to 5.65 times (**RQ5.5**). In tasks 5 to 8 where there is forcible selection by software as compared with tasks 1 to 4, the mean square error E becomes large. That is, as the degree of uncertainty of information increases, the estimation error increases for both Bayesian inference and BIB inference (**RQ5.6**). The average value of the mean squared error excluding tasks 1 and 5 and the median of all tasks were smaller in BIB inference than in Bayesian inference. However, these significant differences (T test p <0.05) were not observed except task 8 (**RQ5.7**). **Figure 40** shows the difference in error between Bayesian estimation and BIB estimation. There is no significant difference, but as the degree of uncertainty is higher, the result of BIB inference tends to be slightly advantageous (**RQ5.8**).

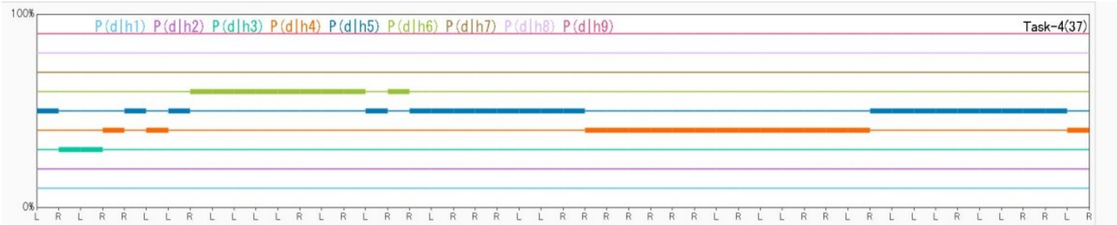


(a)

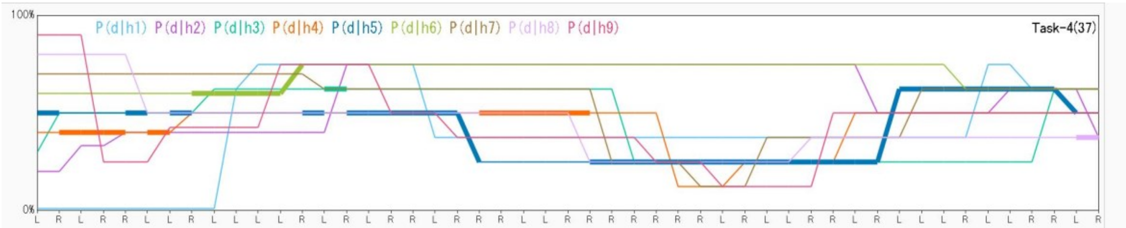


(b)

Figure 35 Change in prior probability $P^t(h_i)$ by Bayesian inference (a) and BIB inference (b). In this graph, one of the results of task 4 is used. The moving-average interval used by inverse Bayesian inference was set to 7. (Also in Figure 36, Figure 38, and Figure 41)

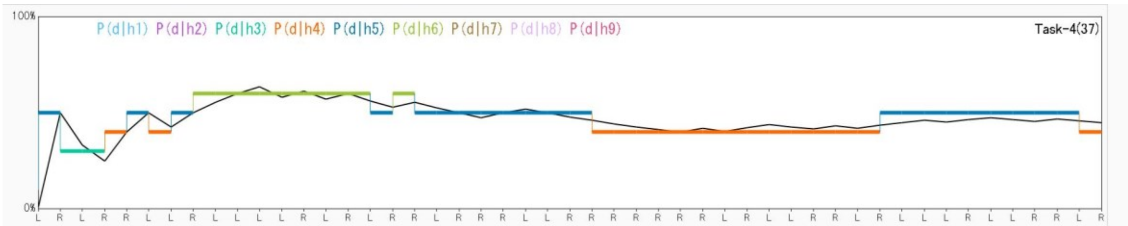


(a)

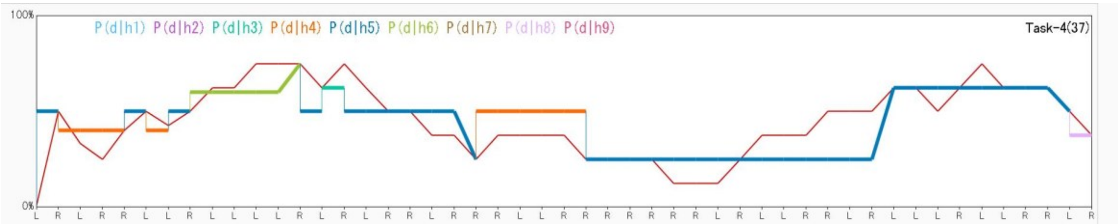


(b)

Figure 36 Change in likelihood $P^t(d|h_i)$ of each hypothesis of Bayesian inference (a) and BIB inference (b). The thick segments on these plots indicate the hypothesis having the maximum prior probability $P^t(h_i)$. Likelihood is constant in Bayesian inference, but the likelihood of BIB inference is updated according to the moving average $P^t_{mov}(d)$.



(a)



)

(b)

Figure 37 (a) Change in cumulative average from the start position P^t_{ac}(d) (black) and likelihood $P^t(d|h_i)$ for maximum prior probability $P^t(h_i)$ by Bayesian inference. (b) Moving average $P^t_{mov}(d)$ (red) and likelihood by BIB inference. The likelihood by Bayesian inference change in line with the average from start time while those by BIB inference change in line with the moving average.

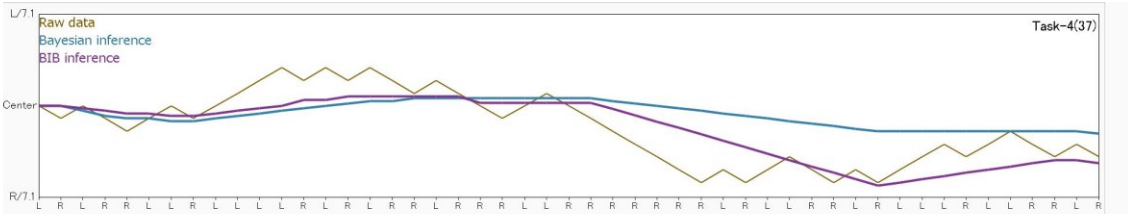


Figure 38 Comparison of actual measurements and inferred values at each step. The yellow plot shows cumulative change of experimental data $c_{raw}{}^t$ and the turquoise and purple plots show cumulative change $c_{inf}{}^t$ of the results obtained by Bayesian and BIB inference, respectively. Bayesian inference cannot keep up with sudden changes, but BIB inference can.

Table 22 Mean square error of inferred values $c_{inf}{}^t$ relative to experimental data $c_{raw}{}^t$ for each task.

	task1		task2		task3		task4		task5		task6		task7		task8	
	Bayes	BIB	Bayes	BIB	Bayes	BIB	Bayes	BIB	Bayes	BIB	Bayes	BIB	Bayes	BIB	Bayes	BIB
mean (E)	0.72	0.95	0.76	0.67	0.64	0.64	8.13	6.73	30.52	31.91	23.25	18.57	19.03	15.75	18.76	14.76
$E^{1/2}$	0.85	0.98	0.87	0.82	0.80	0.80	2.85	2.60	5.52	5.65	4.82	4.31	4.36	3.97	4.33	3.84
SD	0.54	2.26	0.45	0.37	0.46	0.60	6.76	4.68	22.66	35.93	25.89	25.16	20.05	15.10	18.57	14.79
median	0.54	0.52	0.57	0.55	0.53	0.51	5.64	5.54	26.96	18.42	16.51	13.10	13.51	13.34	12.07	10.67
T test p	4.4E-01		8.1E-02		9.9E-01		9.7E-02		6.8E-01		1.9E-01		5.9E-02		2.8E-03	

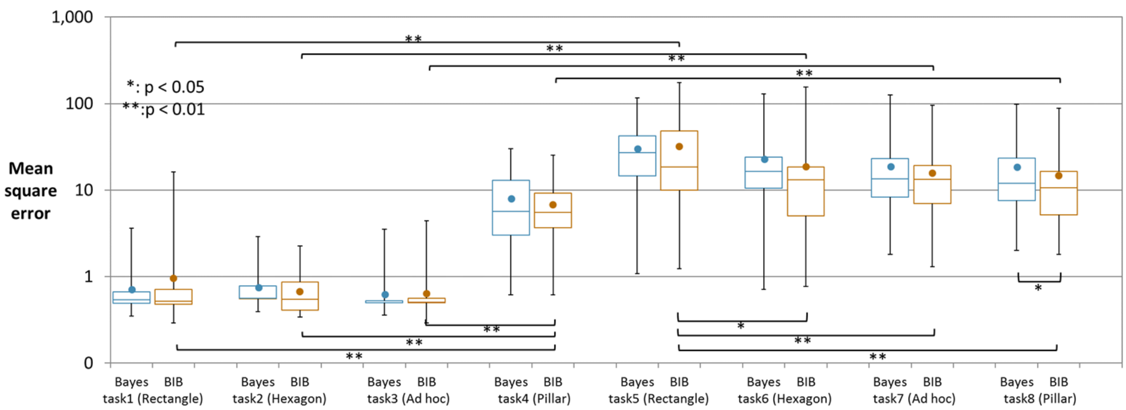


Figure 39 Mean square error of inferred values $c_{inf}{}^t$ relative to experimental data $c_{raw}{}^t$ for each task.

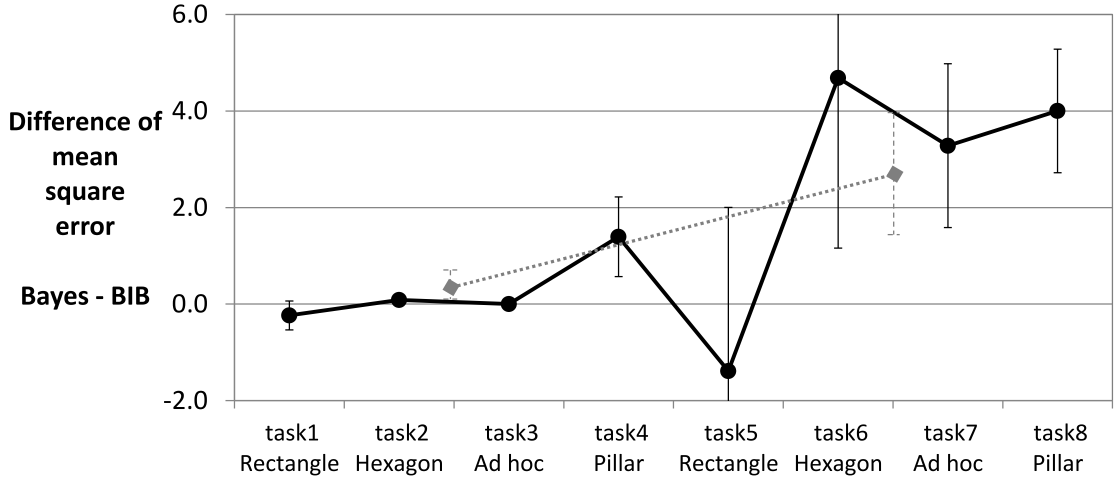


Figure 40 Difference in mean squared error between Bayesian inference and BIB inference. There is no significant difference, but as the degree of uncertainty is higher, the result of BIB inference tends to be slightly advantageous. The dotted line shows the average value of task 1-4 and task 5-8.

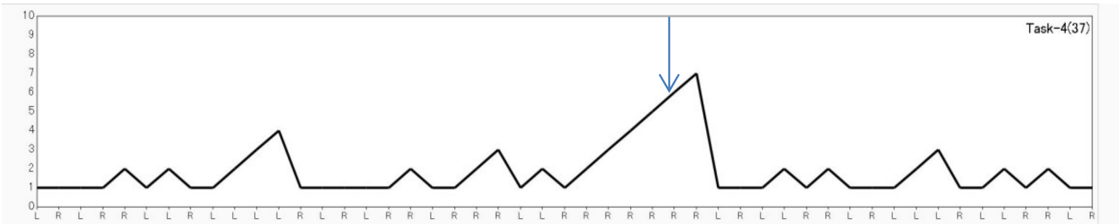
5.4 BIB inference and moving average

As shown in **Section 5.3** , the goal position, the judgment time, and percentages of the unreasonable decisions did not change depending on the presence or absence of software intervention (**RQ5.4**), but a large change was observed in the inference error (**RQ5.6**). Also, when there was software intervention, the error of BIB inference was somewhat lower than the Bayesian inference (**RQ5.8**). In the event that data behavior suddenly changes, a divergence arises between the cumulative average $P^t_{ac}(d)$ and the moving average $P^t_{mov}(d)$. For example, when suddenly selecting the same direction in succession in contrast to past behavior as shown by the arrow in **Figure 41 (a)**, a large divergence arises between the cumulative average $P^t_{ac}(d)$ and the moving average $P^t_{mov}(d)$ as shown in **Figure 41 (b)** in black and colored lines, respectively. As shown in **Figure 42**, such a divergence increases as the degree of uncertainty increases, and shows the same tendency as the inference error for each task. In Bayesian inference, the prior probability $P^t(h_i)$ of the hypothesis with the likelihood $P^t(d|h_i)$ near the cumulative average $P^t_{ac}(d)$ is high, but in BIB inference, the likelihood $P^t(d|h_i)$ is updated by the value of moving average $P^t_{mov}(d)$, and as a result, the prior probability $P^t(h_i)$ of that hypothesis becomes high.

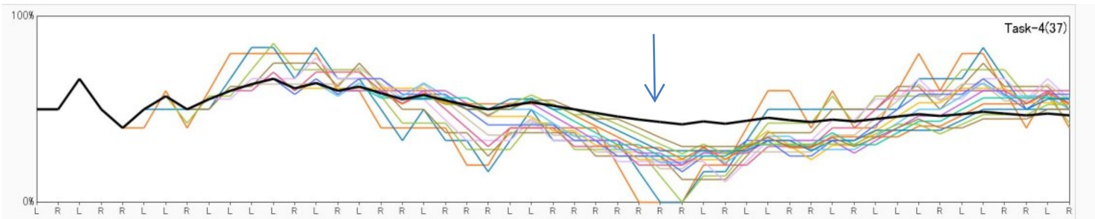
The results of an experiment involving movement through the first person's perspective revealed many judgments that could be considered unreasonable when viewed in objective terms. In addition, judgments of this type would suddenly appear. This is thought to be because the subject would suddenly feel that "I went far" and think "I must return." We modeled such a decision-making process by both Bayesian inference and BIB inference and showed that the latter could deal with such sudden changes.

As Gigerenzer [22], Knill [12], and Manktelow [24] pointed out, the results of this experiment also suggest that human decision making process can be modeled by Bayesian inference. However, as the degree of uncertainty of information increased, accuracy of inference tended to decrease. BIB inference can also be estimated with the same precision as Bayesian inference, and there is a possibility of being able to estimate with higher accuracy in situations where the behavior changes suddenly.

In the BIB inference mentioned in **Section 5.2** , moving average was applied to the likelihood in the process of inverse Bayesian inference. In addition to this, it is also possible to use weighted averages, exponential moving averages and others. The interval of moving average corresponds to how much information of the past is taken as an internal state. This value was about 7 seconds in the experiment described in **Section 5.4** . We will consider the optimal method to capture from the past information into the internal state.



(a)



(b)

Figure 41 Change in number of times same direction was consecutively selected (a) and change in cumulative average $P^t_{ac}(d)$ (black) and (b) moving average $P^t_{mov}(d)$ (colored, interval: 4 – 20). As indicated by the arrow, consecutive selection of the right direction results in a divergence between cumulative average and moving average. A difference consequently arises between the results of Bayesian inference and BIB inference with the latter performing sensitive tracking of immediately previous changes.

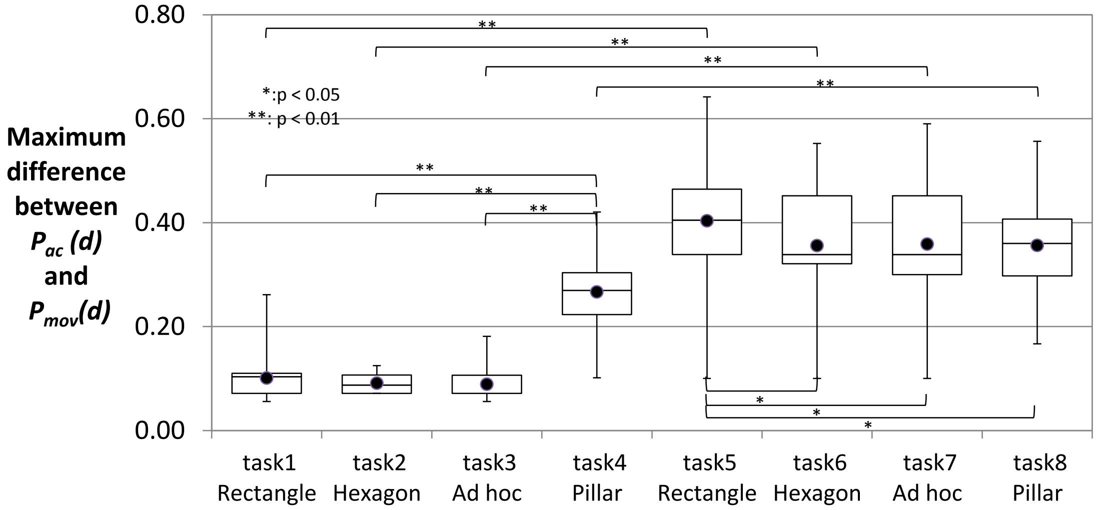


Figure 42 Maximum difference between the cumulative average $P^{t}_{ac}(d)$ and the moving average $P^{t}_{mov}(d)$.

5.5 Brain that prevents straight walking

In the experiments of moving in a virtual three-dimensional space, many unreasonable judgments were observed. Depending on the three-dimensional model, the goal position, judgment time, and unreasonable judgment ratio changed. However, differences due to changes in the degree of uncertainty were not observed. The change in the direction is more influential than the change in the distance, and the angle of turning also affects. As a result of applying the time series data of the judgment of the subjects to the Bayesian and BIB inference, both results are roughly consistent with the actual data. As the degree of uncertainty increases, the inference error also increases. The result of BIB inference tends to match the actual data rather than that of the Bayesian inference. It is thought that the decision-making process in which the strategy suddenly changes can be modeled by BIB inference. In the future, we are planning experiments to verify whether we can walk straight in real city blocks.

Chapter 6 Decision making process of groups: Conformity and repulsion affected by surrounding opinions

Summary of Chapter 6

In this chapter, modeling decision-making of group will be described. Conforming with surrounding opinions is called a bandwagon effect, and repulsing is called an underdog effect. The bandwagon effect can be said the shrinking process of opinions, and the underdog effect can be said the expansion process of opinion. These effects based on poll results at the time of election have been formulated for a long time and studied through virtual voting experiments. It is also known that the same effect have been recognized in financial markets and enterprises' technology introduction. In this chapter, I examined the transition data of four companies' polls during the 2016 US presidential election. There can be a bandwagon effect if there is a positive correlation between the support rate of the previous survey result and the change of the current support rate, and if there is a negative correlation there is an underdog effect. As a result of the verification, a negative correlation showing an underdog effect was observed. The underdog effect tends to be close battle, in fact, the support rate of both candidates had been around 50% until the end. In addition, according to the support political party, there was a stronger tendency of the underdog effect in the group without the supportive party. The mean squared error between the estimated value by the linear regression model and the actual transition was 1.10% on average for the four companies. In addition, the mean square error of the model by Bayesian inference and *inverse Bayesian inference* are 0.40% and 0.24%, respectively, and the *inverse Bayesian inference* can be said to be the most effective model. This is probably because Bayesian inference has a strong tendency to result in the bandwagon effect, so it was difficult to adapt to this data. On the other hand, *inverse Bayesian inference* is considered to be applicable to both effects. Furthermore, it can also be applied to economic indicators such as stock price and exchange rate, which is another example of group decision making.

6.1 Bandwagon and under dog effects

When making a decision, a bandwagon effect synchronizing with the surroundings and an underdog effect which rebounds to the surroundings are known. These effects are confirmed by the influence of opinion polls before the election [51], financial transactions such as stocks and currency exchange [52], decision making at the time of introducing new technology of the company [53], and so on. Also, psychological experiments have verified the influence of majority factions on minorities [54]. On the other hand, Bayesian inference is widely used for modeling human decision making [22]. There are also attempts to apply Bayesian inference to model the bandwagon effect [73]. However, psychological experiments sometimes result in the opposite result of Bayesian inference [39], and modeling of such *Anti-Bayesian* is also attempted [40]. Apart from Anti-Bayesian, *inverse Bayesian inference* that can explain various events has been proposed [41] [1]. Inverse Bayesian inference is effective for modeling human decision making and transition of economic indicators [120]. The purpose of this chapter is modeling of voting behavior that influences the results of public opinion survey using Bayesian inference and inverse Bayesian inference.

Regarding the bandwagon effect and the underdog effect in elections, Simon formulated the relationship between opinion poll prediction and actual voting behavior [55]. In addition, many virtual voting experiments to verify these effects have been done, and the bandwagon effect has been confirmed [56] [57] [58] [59] [60]. *Pivotal voter model* based on the hypothesis that voters make voluntary and rational choices has been proposed. In this model, the expected value is calculated from the cost for voting, the degree of influence given by the vote, and the profit from the voting result, and whether or not to vote is judged by this expected value [61]. On the other hand, the *strategic elites model* focuses on behaviors such as candidates, political parties, and leaders of profit organizations [62]. In recent years, with the widespread use of SNS, the information obtained by voters is increasing. The results of public opinion polls are said to affect the voting behavior, and some countries restrict public disclosure during the election period [51]. Also, not only in the US presidential elections in 2000 and 2016, but also in other countries, the case of *close battle* in important elections is increasing.

Nurkse [63] had earlier warned about the dangers of the bandwagon effect that would cause instability in financial markets [52]. Allen et al. analyzed the foreign exchange trends forecasted by the *chartists* who make decisions based on past experience, not economic theory [64]. When the bandwagon effect is seen, the elasticity of expectations for exchange rate fluctuation increases and the market becomes unstable. Siddiqi showed that the broker can control equilibrium state of stock price (perfect Bayesian equilibrium) in Bayesian game by three people in institutional investors, individual investors and brokers in the stock market [65] [66].

Even when companies adopt new technology, it is known that the bandwagon effect can be seen [53]. Decision making of technology introduction is influenced not only by reasonable judgment such as technical or profitability (i.e. efficient-choice theories) but also by the number and reputation that has been adopted in advance (i.e. fad theories) [67] [68].

Many psychological studies have examined the tendency of humans to synchronize with surrounding situations. Asch conducted an experiment to answer the length of the line segment drawn on the card in a group with seven confederates for each subject [54]. All seven confederates return the same wrong answer, and finally the subject answers. It has been shown that at least 75% tend to synchronize with confederates' answers at least once and tend to be in line with the majority opinion. Moscovici, on the other hand, examined the influence of the minority on the majority [69]. In the experiment, four subjects and two confederates answer the color of the projected slide. Even if the minority's confederates are wrong, when the answers are consistent, it has become clear that they affect the majority. Moreover, the influence tends to be strong when the way of thinking is flexible [70] [71] and the social attributes are similar [72].

As mentioned below, the *Bayesian inference* has become widely used to model various cognitive and decision-making processes and hundreds of studies have been done. Gigerenzer showed by experiment that human thinking is performed along the lines of Bayesian inference [22]. Meanwhile, Knill described how humans and other living organisms perform stochastic inference based on uncertain information obtained from the real world and reported that many experimentally observed examples could be explained by a Bayesian perceptual system [23]. In contrast to reacting suddenly to particular input, this system integrates information that propagates over space and time. In addition, Manktelow classified and analyzed how humans perceive a variety of uncertain events such as coin-toss games, probability that a weather report is accurate, existence of the Loch Ness Monster, etc. It was described here that Bayesian inference, which is used to update a person's hypothesis after observing an event as posterior probability, is the basis for human intuitive cognition [24]. There are also many studies that children's developmental process is in line with Bayesian principles [25], [26], [27], [28]. Pellicano, meanwhile, described how the perceptual process of autistic patients follows Bayesian inference and that autistic people tend to dislike unconventional stimuli that do not agree with experience [29]. This result suggests the ability and limitations of the Bayesian inference.

There are a few studies that associate the bandwagon effect with Bayesian inference. Rosenkopf et al. pointed out that the bandwagon effect was observed when adopting new technology, the criteria were updated according to the circumstances, and that the process was a Bayesian mechanism [53]. Chappell et al. proposed a *Bayesian Bandwagon Model* [73]. They analyzed data on the decision process of the federal funds rate by members of the Federal Open Market Committee (FOMC) of the Federal Reserve Board (FRB). However, the bandwagon effect was not seen in FOMC.

There are also some reports on the phenomenon opposite to Bayesian inference (i.e. *anti-Bayesian*). Brayanov et al. have found that classical *size-weight illusion* (or Charpentier-Koseleff illusion) has been recognized by experiments as opposed to Bayesian inference [39]. Wei et al. proposed a model to derive the mechanism of *anti-Bayesian* discovered by Brayanov et al [40]. By making the likelihood asymmetric by sensory noise inside the observer, contrary to Bayesian inference, the bias of repulsion acts on the prior probability distribution. Thomas et al. Proposed a classification method of patterns by anti-Bayesian approach using moments of *order statistics* [123] [124] [125]. Accurate classification is possible with a small number of samples away from the average in the parameter space.

The aim of this chapter is modeling of voting behavior in which bandwagon effect and underdog effect appear using Bayesian inference and inverse Bayesian inference. For each public opinion research company, we quantitatively evaluate the bandwagon and underdog effect with the presence or absence of a supporting party. Moreover, we verify whether the bandwagon effect and the underdog effect can be modeled by Bayesian inference and further verify by using inverse Bayesian inference.

6.2 Materials and methods

6.2.1 Data

Of the opinion polls during the presidential election of the US in 2016, we used the data of companies (Rasmussen, CNN, FOX, and Ipsos / Reuters) surveyed by supporting political parties [126]. **Table 23** shows the number of surveys, the method of survey, and the average of the number of observers in surveys conducted by each company. Only answers that support Trump candidates and Clinton candidates are used, and responses that support unanswered or other candidates have been deleted. **Figure 43** shows an example of data. In addition, the time when the public opinion survey was carried out is not equally spaced, but it was analyzed as time series data of the support rate every time.

Table 23 The number of surveys, the method of survey, and the average of the number of observers in surveys conducted by each company.

	all	democrat	republican	independent	Mode	Number of Observations
Rasmussen	33	12	12	12	IVR/Online	1152
CNN	18	18	18	18	Live Phone	909
FOX	22	16	16	16	Live Phone	1010
Ipsos/Reuters	41	26	26	26	Internet	1319

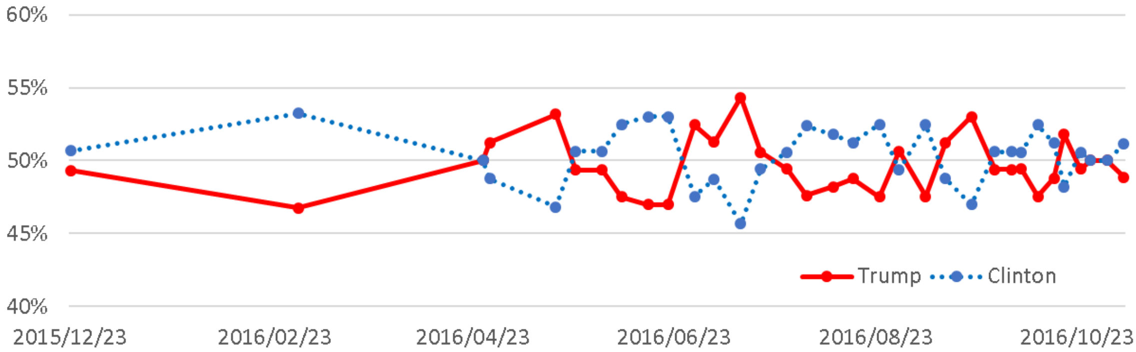


Figure 43 Changes in the support rate by public opinion survey. The research company is Rasmussen, and the subjects of the survey are all the voters including the presence or absence of a supporting party.

6.2.2 Research questions

The research questions are as follows:

RQ6.1 Does the bandwagon effect or the underdog effect be recognized from the trend of poll results?

RQ6.2 Do these effects differ depending on the presence or absence of a supporting party?

RQ6.3 Is Bayesian inference effective for modeling transition of public opinion survey?

RQ6.4 Is inverse Bayesian inference effective for modeling transition of public opinion survey?

RQ6.5 Which model is most effective for modeling the transition of public opinion survey?

6.2.3 Methods

As the first model, suppose that the support rate further increases when the support rate is high, and when it decreases when it is low, there is a bandwagon effect (model 1). Let T be the number of surveys. Let $R(t)$ be the support rate of the Trump candidate for the t (= 1, 2, .., T) th survey, and let $d(t) = R(t + 1) - R(t)$ be the change amount of the support rate from the t th time to the t+1 th time. Here, if there is a bandwagon effect, the correlation coefficient between $R(t)$ and $d(t)$ has a positive value, and if there is an underdog effect, it should be a negative value. Equation (14) is a linear regression of this model, and equation (15) is a model function when $\varepsilon(t) = 0$.

$$d(t) = a_1 * R(t) + b_1 + \varepsilon(t) \quad (14)$$

$$M_1(t+1) = M_1(t) + a_1 * R(t) + b_1 \quad (15)$$

Apart from this, it can also be said that there was a bandwagon effect when the support rate further increases when the support rate increases and decreases when it decreases (model 2). At this time, if the correlation coefficient between $d(t - 1)$ and $d(t)$ is a positive value, there is a bandwagon effect, and if it is a negative value, it can be said that there was an underdog effect. Equations (16) and (17) show approximate equation and model function, respectively.

$$d(t) = a_2 * d(t\text{-}1) + b_2 + \varepsilon(t) \quad (16)$$

$$M_2(t+1) = M_2(t) + a_2 * d(t\text{-}1) + b_2 \quad (17)$$

Figure 44 shows the relationship between correlation coefficient and these effects.

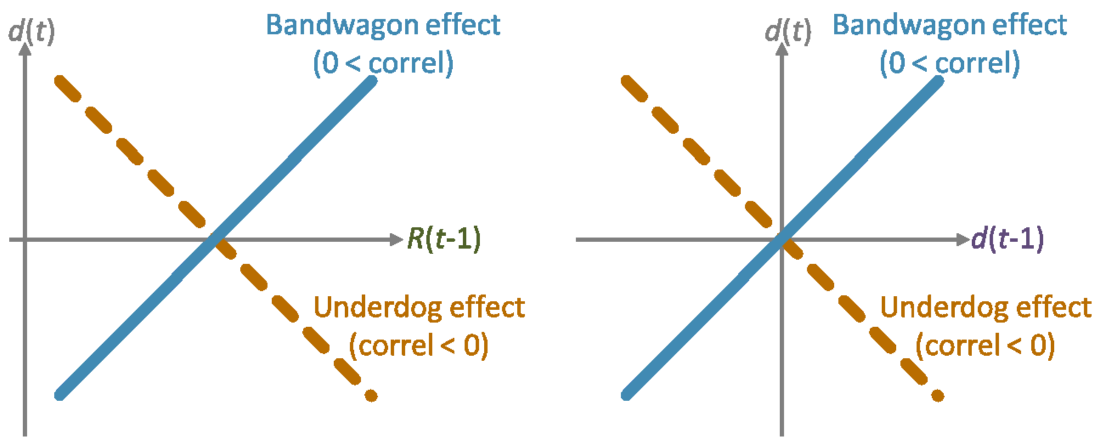


Figure 44 The sign of the correlation coefficient determines the existence of the bandwagon effect or the underdog effect. (a) model 1 ($d(t)$ vs $R(t)$) and (b) model 2 ($d(t)$ vs. $d(t\text{-}1)$).

In Bayesian inference, posterior probability is calculated from the prior probability and likelihood for the data sequence, and in the next step, the prior probability is replaced with the posterior probability and the calculation is advanced. In the inverse Bayesian inference, the likelihood is further updated using the information of the past history. A method combining Bayesian inference and inverse Bayesian inference is called BIB inference [1]. Details of these calculation methods are described in the Appendix. Let h^* be the hypothesis with the highest prior probability $P^t(h^*)$ at t-th step. In this hypothesis, let probability $P^t(d_u|h^*)$ be the likelihood that the amount of change in value falls in the interval from d_u to d_{u+1}. The model functions obtained by Bayesian inference $M_b(t)$ and BIB inference $M_{bib}(t)$ are described in **Chapter 3** as equations (12) and (13).

6.3 Results

6.3.1 Bandwagon or underdog effect?

Figure 45 (a) shows the relationship between $R(t)$ and $d(t)$ of model 1, and (b) shows the relationship between $d(t - 1)$ and $d(t)$ of model 2. In both models, negative correlations were seen, and it can be said that there was an underdog effect (**RQ6.1**).

Table 24 shows the slope (a_1, a_2), intercept (b_1, b_2), and correlation coefficient of the approximate expression using data from each survey company. In any of the data, the correlation of model 2 tends to be lower than that of model 1. **Figure 46** shows the results estimated from the model equations (16) and (18) using the coefficients of the obtained regression equation.

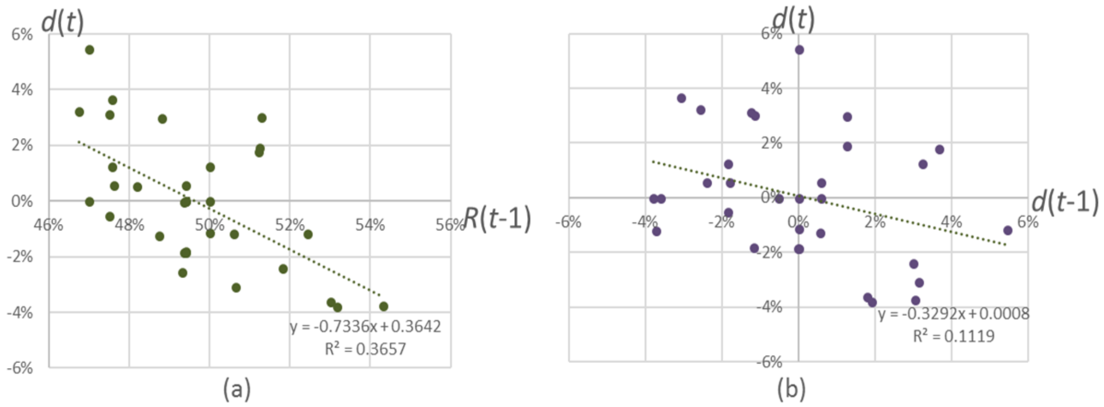


Figure 45 The relationship between support rate $R(t)$ and change of the support rate $d(t)$ of model 1(a), and the relationship between $d(t - 1)$ and $d(t)$ of model 2(b). There was an underdog effect.

Table 24 The slope, intercept and correlation coefficient of the approximate expression using data from each survey company.

	Model 1			Model 2		
	a_1	b_1	Correl.	a_2	b_2	Correl.
Rasmussen	-0.734	0.364	-0.605	-0.329	0.001	-0.335
CNN	-0.727	0.344	-0.720	-0.346	0.006	-0.359
FOX	-0.999	0.479	-0.765	-0.294	0.002	-0.317
Ipsos/Reuters	-0.604	0.276	-0.582	-0.453	0.001	-0.475

Figure 46 The results estimated from the model equations (15) and (17) using the coefficients of the obtained regression equation.

6.3.2 Influence of existence of supporting party

Next, with respect to **RQ6.2**, we analyzed the data of the group that supports the Democratic Party, the Republican Party, and the group without the supportive party. Plotted in the same way as in **Figure 45** is shown in **Figure 47**. Compared with the case with a supportive party, data without supportive parties has large variance of both the support rate $R(t)$ and the change amount $d(t)$. Both have negative correlations, which are underdog effects. **Table 25** shows the slope, intercept, and correlation coefficient of the approximate expression from data without supporting party. Compared to the data of the case with the whole or supporting political party, the data without the supporting party has a high correlation coefficient.

Figure 47 The relationship between support rate *R*(*t*) and change of the support rate *d*(*t*) of model 1(a), and the relationship between *d*(*t* - 1) and *d*(*t*) of model 2(b) for the group of Republican support, Democrat support, and independent. The research company is Ipsos / Reuters.

Table 25 The slope, intercept and correlation coefficient of the approximate expression using data from each survey company for the group of independent.

	Model 1			Model 2		
	a_1	b_1	Correl.	a_2	b_2	Correl.
Rasmussen	-1.244	0.693	-0.833	-0.596	0.000	-0.600
CNN	-0.820	0.395	-0.690	-0.251	0.008	-0.251
FOX	-1.140	0.049	-0.765	-0.735	0.009	-0.717
Ipsos/Reuters	-0.963	0.549	-0.691	-0.480	-0.013	-0.481

6.3.3 Bayesian and inverse Bayesian inference

Figure 48 shows the results estimated by the model functions using Bayesian inference $M_b(t)$ and BIB inference $M_{bib}(t)$. **Figure 49** shows the relationship between the support rate (model 1) and the change amount of the previous support rate (model 2) with respect to the change amount of the estimated value by the Bayesian inference ($dM_b(t) = M_b(t+1) - M_b(t)$). In both cases, a positive correlation was observed, indicating the behavior of the bandwagon effect. Especially the correlation coefficient of model 2 is high. On the other hand, as shown in **Figure 50**, this property is relaxed in the model using BIB inference ($dM_{bib}(t) = M_{bib}(t+1) - M_{bib}(t)$). Here also positive correlation was observed, but the trend is weak.

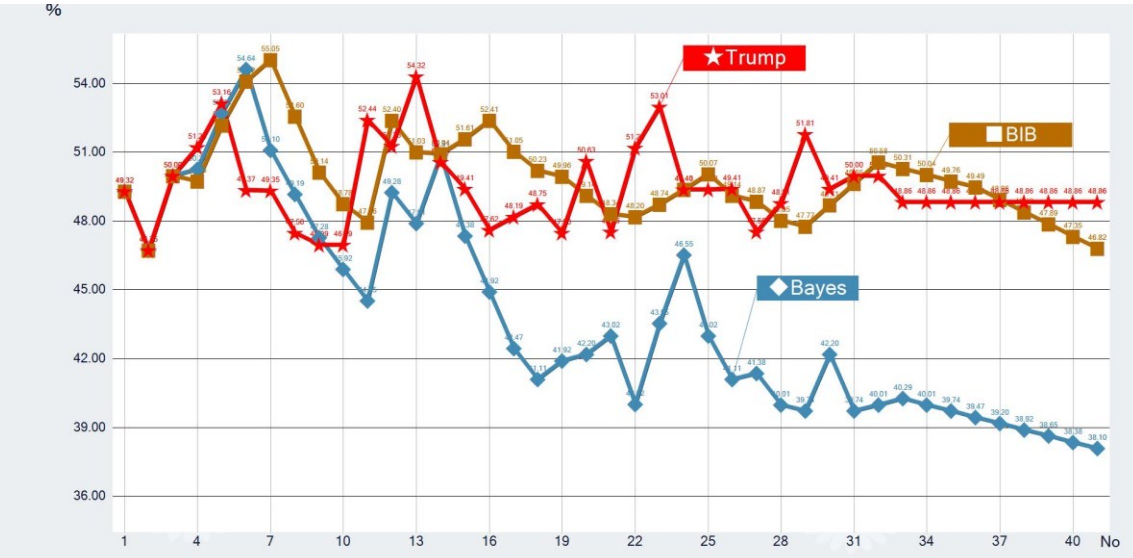


Figure 48 Estimation results by model using Bayesian inference ($M_b(t)$) and BIB inference ($M_{bib}(t)$).

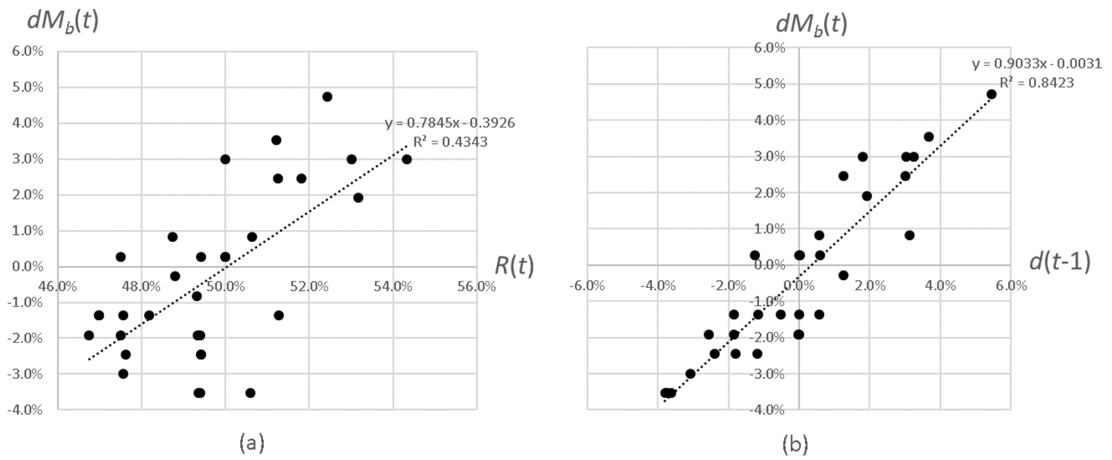


Figure 49 Relationship between (a) change amount (*$dM_b(t)$*) and support rate(*$R(t)$*), (b) change amount (*$dM_b(t)$*) and change amount of previous support rate (*$d(t$*-1)), as a result of model using Bayesian inference.

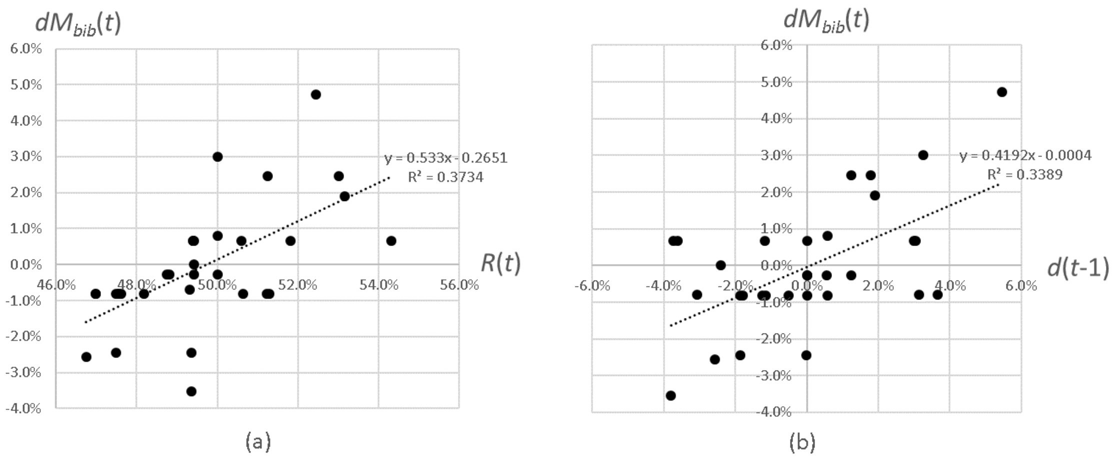


Figure 50 Relationship between (a) change amount (*$dM_{bib}(t)$*) and support rate(*$R(t)$*), (b) change amount (*$dM_{bib}(t)$*) and change amount of previous support rate (*$d(t$*-1)), as a result of model using BIB inference.

6.3.4 Evaluation

For models 1 and 2, coefficients were calculated by the least squares method using all data. Therefore, coefficients should be calculated sequentially for each step and the model function (equations (16) and (18)) should be updated. **Figure 51** shows the estimation results by the updated model functions. **Figure 52** shows the mean square errors between the estimation results of four models ($M_1(t)$, $M_2(t)$, $M_b(t)$, $M_{bib}(t)$) and the original data $R(t)$. Estimation errors by model using Bayesian inference and BIB inference are small compared to those of models 1 and 2

(**RQ6.3** and **RQ6.4**). Furthermore, the error of the model using BIB inference tends to be smaller than the model of Bayesian inference. Therefore, it can be said that the model using the BIB inference most expresses the original data (**RQ6.5**).

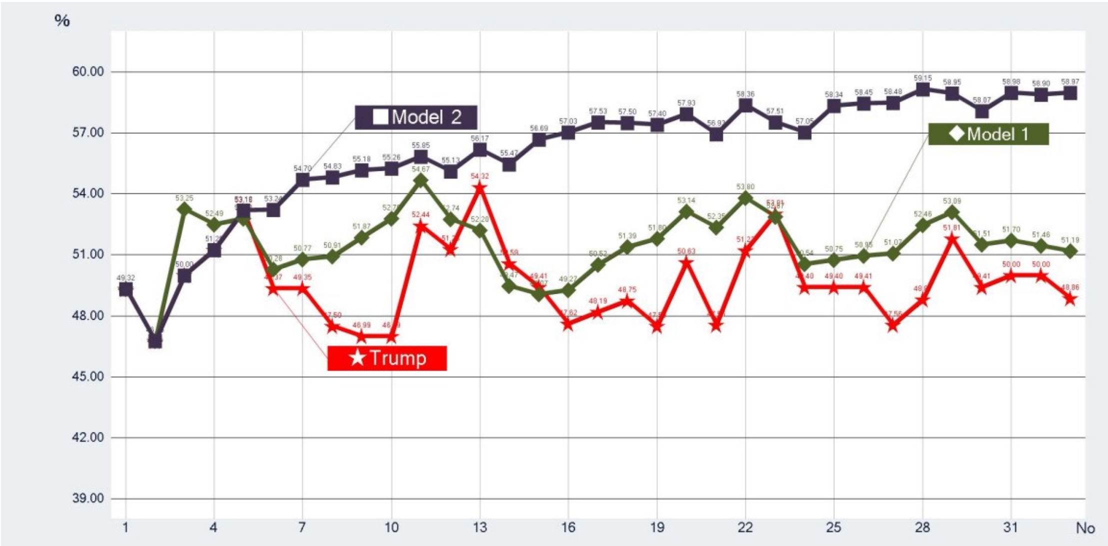


Figure 51 Estimation results when coefficients a1, b1, a2, b2 of models 1 and 2 by regression are updated step by step.

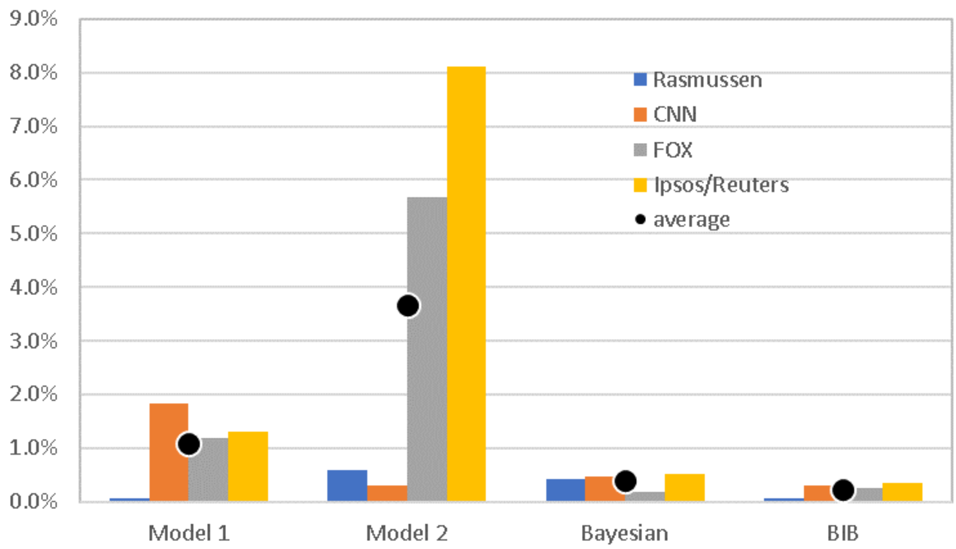
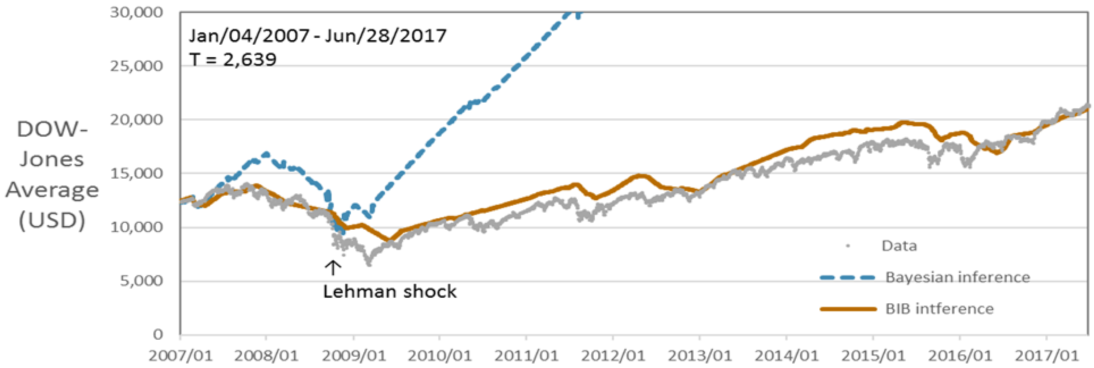


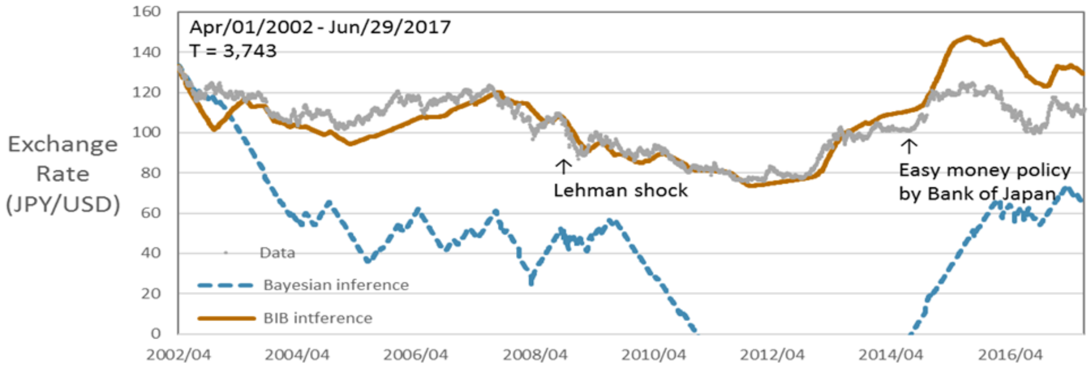
Figure 52 Mean square error of the original data and estimation results of the four models.

6.4 Application to economic indicators

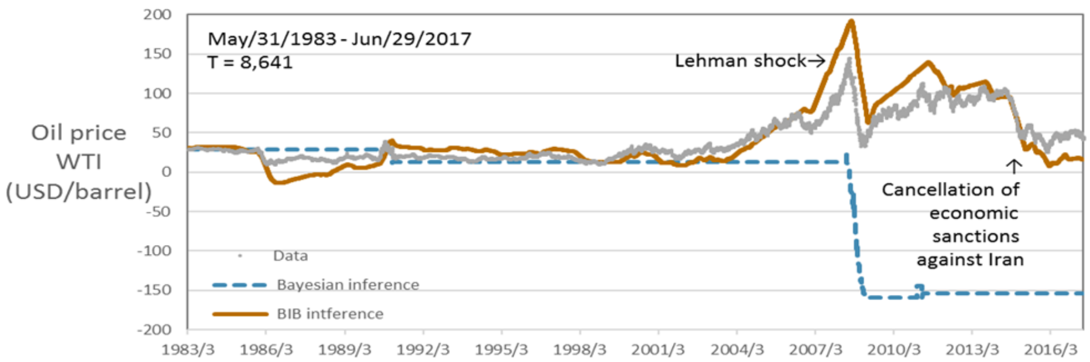
The inference examples of economic indicators, such as stock price, exchange rate, and oil price are shown in **Figure 53**. These indicators also represent the *culmination* of many human decision-making, and sometimes they show unexpected behavior. As shown in the figures, when the number of data is large, the model using BIB inference can better represent the actual data.



(a)



(b)



(c)

Figure 53 Inference examples of economic indicators. (a) Stock price [127], (b) Exchange rate of Japanese Yen vs US Dollar [128], and (c) Oil price of West Texas Intermediate [129]. The number of hypotheses *N* is 28, the number of sections of values *M* is 28, and the interval of

moving average is 32. Even in unconventional situations such as the Lehman shock, BIB inference can follow the actual data.

6.5 The limitation of Bayesian inference and the possibility of inverse Bayesian inference

In the US presidential election in 2016, the change in the support rate had a negative correlation with the "support rate" and "the change in the last support rate", and the tendency was stronger in the former. That is, it can be said that there was an underdog effect. When there is an underdog effect, it becomes close battle. Also, the underdog effect was more prominent in groups without a supporting party. When voting behavior was modeled using Bayesian inference, there was a strong bandwagon effect. This is because the tendency to select hypotheses with high prior probability is enhanced. On the other hand, even in the BIB inference, although the bandwagon effect is similarly observed, the absolute value of the correlation coefficient is relatively small. In the model functions using two regression, Bayesian inference and BIB inference, the model function using BIB inference was closest to the original data. This is because the likelihood is updated in the calculation process of the inverse Bayesian inference and balanced.

We modeled the transition of public opinion poll of the US presidential election in 2016 using two regression formulas. As a result, the underdog effect was seen in data of each survey company. This trend was prominent for voters who do not have a supporting party. On the other hand, the model using Bayesian inference could be estimated with higher accuracy than the regression models. However, this model tends to lead to the bandwagon effect. Furthermore, in the model using BIB inference that combines Bayesian inference and inverse Bayesian inference, the tendency of the bandwagon effect is relaxed and it can be estimated with higher accuracy.

In this consideration, we adopt the assumption that "the voters have made decisions by seeing the results of each survey company." However, in reality they are making decisions by referring to various information. There are various paths for propagation of information, and analysis of propagation of opinion on the network has been studied numerously [130]. In the future, we plan to investigate in various ways, such as the case of other elections, the difference in the public availability of opinion poll results.

Chapter 7 Conclusion: Toward a new harmony of society

7.1 New harmony of society featuring inverse Bayesian inference

In the real world, a large number of decision makers are involved, the uncertainty is high, and various stakeholders make decisions with their respective motives. As shown in Figure 54 (a), mismatch of value criteria creates different rationality and prevents consensus formation. In addition, irrational judgment is frequently done also in individual decision making process. Furthermore, it is influenced to be conformed and repulsed against surrounding opinions.

In **Chapter 4** to **Chapter 6**, I examined the nature of these human decision making and consensus building, and tried to model. The models of individual decision making process including irrational judgment and the model of group decision process affected by surrounding opinion are effective methods. As shown in Figure 54 (b), by incorporating these models, I will establish a methodology for decision making and consensus building that brings new harmony of society. In order to realize this, the flow of information is divided into a logical layer and an uncertain layer as shown in Figure 54 (c). In the logical layer, it is based on the description language and the visualization / analysis tool described in **Chapter 4**. This layer receives any information from the uncertain layer and simulates the whole behavior with strict calculation. Uncertain layer consists of individual decision making models and influence models of opinions of each other. In the individual decision making model described in **Chapter 5**, even if it is the result of the decision of the person him / herself, estimate whether it is objectively reasonable or unreasonable. Furthermore, influence models described in **Chapter 6** estimate the influence from others and the impact on others. However, it would be difficult to express them with reliable numerical values. In reality, it is thought that probabilistic process will be used. Here, it is considered that a methodology of new consensus formation can be established by using *inverse Bayesian inference*.

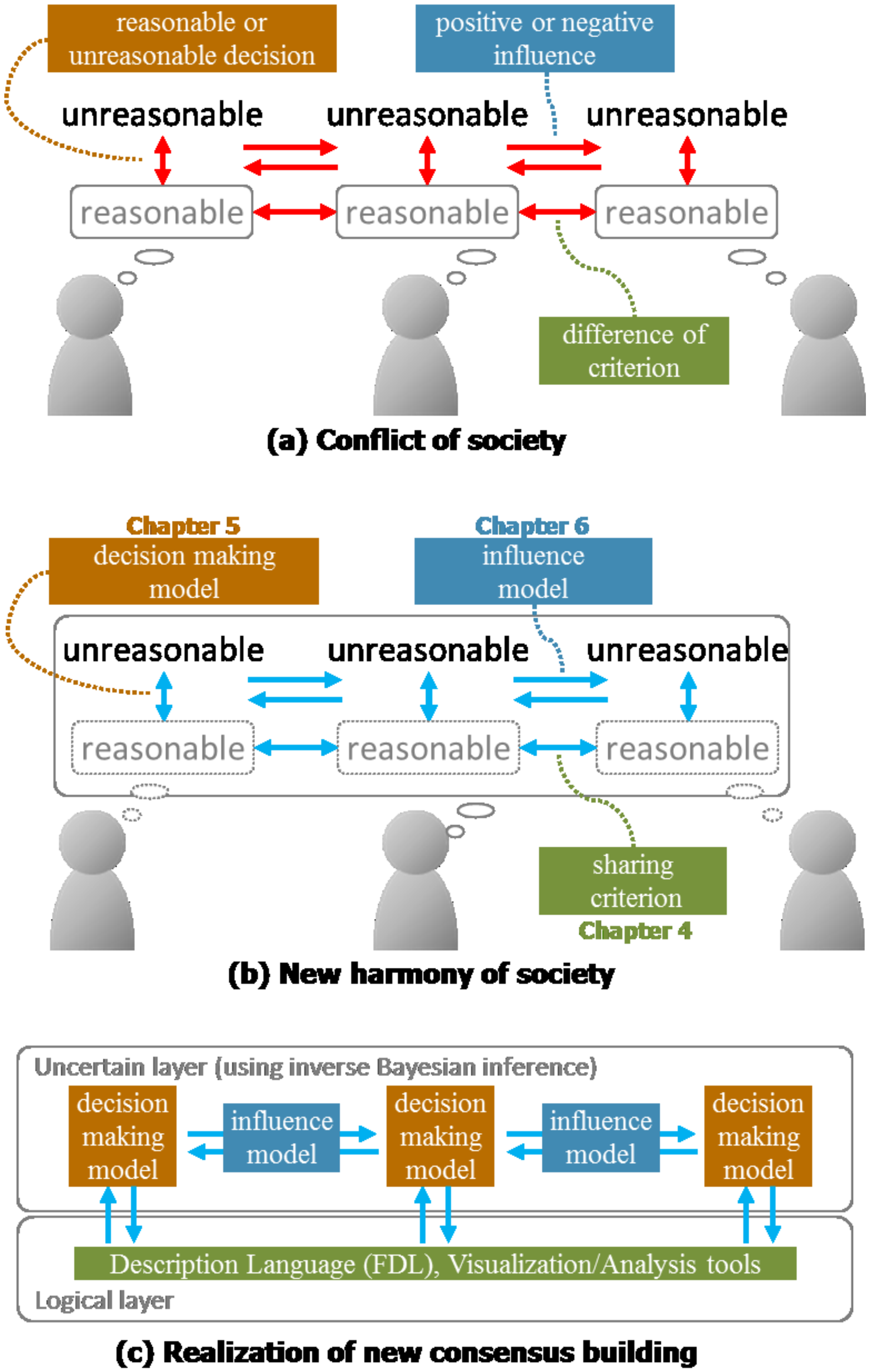


Figure 54 Methodology for making decisions and consensus building to realize new harmony

7.2 Conclusion and future prospects

In this doctoral thesis, I showed that human decision making process can be modeled by *inverse Bayesian inference* with examples. The *inverse Bayesian inference* can be characterized by determining the likelihood from the past history.

In **Chapter 4**, the methodology for conveniently performing these calculations, including *inverse Bayesian inference*, was described. With a newly developed description language, it is possible to perform complicated calculation without programming knowledge. The shrinking process towards the fixation of the logic to solve the problem and the expansion process to fluidization are repeated. In the shrinking process, fixation of logic is attempted while including uncertain elements. Sensitivity analysis of the fixed logic shows the effect of uncertainty on the whole, and in the expansion process, new elements making up the logic are explored.

In **Chapter 5**, I confirmed the effectiveness of *inverse Bayesian inference* by data on the decision of right and left choices when traveling through the maze on the virtual space. In an environment with high uncertainty, humans frequently make unreasonable judgments objectively. For example, human beings get lost their direction because they have evaluated the distance they have traced so far, either too much or too little. There are a shrinking process to select the optimal solution and an expanding process to search for a different methodology from the optimum solution up to that point. Through the experiments in the maze of virtual space, these shrinking and expanding processes were observed and modeled using *inverse Bayesian inference*.

In **Chapter 6**, I also confirmed the effectiveness of *inverse Bayesian inference* by data on the transition of public opinion survey which is a collection of decision making in the election. The opinions of people are often greatly influenced by the opinions of the surroundings. There are aggregation and divergence of opinions by conformity and repulsion with the surroundings. Aggregation of opinions by conforming with the surroundings is a shrinking process, and divergence of opinion by repulsion is a expanding process. These processes were verified by using polls during the election and economic indicators such as stock price and exchange rate and showed that they can be modeled by *inverse Bayesian inference* as well.

Finally, previous section of this chapter describes the concept for new harmony of society. In order to form an appropriate consensus formation, it was necessary to return to the understanding of the human decision making process itself. The models of individual decision making process including irrational judgment (**Chapter 5**) and the model of group decision process affected by surrounding opinion (**Chapter 6**) will be effective methods. The future plan for building methodology of new harmony of society was described.

Some technical problems such as how long should be traced back and how many hypotheses on the data are appropriate remains. However, it can be a powerful tool in order to build a model in which humans and organisms make decisions in a highly uncertain environment. In the future, I will develop individual trends on infrastructure and modeling decisions as a group and plan to develop a methodology for planning both social hardware and software.

Moreover, the calculation method itself of the *inverse Bayesian inference* is effective for an event which could not be explained by Bayesian inference alone. In particular, it is applicable to many cases where conflicting mechanisms coexist.

References

[1] P.-Y. Gunji, S. Shinohara, T. Haruna and V. Basios, "Inverse Bayesian inference as a key of consciousness featuring a macroscopic quantum logical structure," BioSystems, 152, 44-65, 2016.

[2] Y. Horii, "Chondrule formation: Experimental reproduction by Free-fall," Graduation thesis, Faculty of Science, Kobe University, 1988.

[3] Y. Horii, N. Fujii and H. Takeda, "Hardness analysis of metallic particles in ordinary chondrites," Proceedings of the NIPR Symposium on Antarctic Meteorites 3, 254-263, 1990-10, 1990.

[4] 堀井洋一, 下岡広志, 浅山勇介, 森永速男 , 庫本正, "秋芳洞における石灰華段丘(百枚皿)の形状解析とその成因," 日本洞窟学会第 15 巻 34-41, 1990.

[5] Y. Horry, "What the form tells to us," Master Thesis, Graduate School of Science, Kobe University, 1990.

[6] Y. Horry, "Graphical User Interface for MIDI Signal Generation and Sound Synthesis," Proceedings of the International Computer Music Conference ICMC `94, Aarhus 276-279, https://quod.lib.umich.edu/i/icmc/bbp2372.1994.072/1, 1994.

[7] Y. Horry, K. Anjyo and K. Arai, "Tour into the picture: using a spidery mesh interface to make animation from a single image," Proc. SIGGRAPH '97 (Los Angels, California, August 3-8,1997), 225-232, 1997, http://research.cs.queensu.ca/~jstewart/850/papers/p225-horry.pdf, 1997.

[8] Y. Horry, I. Nakajima, T. Hoshino and Y. Maruyama, "A Passive-style buttonless mobile terminal," IEEE Transactions on Consumer Electronics, Vol. 49, No. 3 530-535, 2003.

[9] E. Mach, "Contributions to the Analysis of the Sensations," Open Court Publishing Co., Chicago, IL USA, 1980.

[10] H. Helmholtz, "The perceptions of Vision (J. P. Southall, Trans.)," Physiological Optics, Vol. III, 1925.

[11] H. Helmholts, In Helmholtz's Treatise on Physiological Optics, Translated from the 3d German edn (Southall, J.P.C., ed.), Dover Publications, 1962.

[12] D. Knill, "Discrimination of planar surface slant from texture: human and ideal observers compared," Vision Res. 38, 1683–1711, 1998.

[13] S. Shimozaki, "An ideal observer with channels versus feature-independent processing of spatial frequency and orientation in visual search performance," J. Opt. Soc. Am. A Opt. Image Sci. Vis. 20, 2197–2215, 2003.

[14] R. Murray, "Saccadic and perceptual. performance in visual search tasks. II. Letter discrimination," J. Opt. Soc. Am. A Opt. Image Sci. Vis. 20, 1356–1370, 2003.

[15] J. Saunders and D. Knill, "Visual feedback control of hand movements," J. Neurosci. 24, 3223–3234, 2004.

[16] M. Hattori, "Adaptive Heuristics of Covariation Detection: A Model of Causal Induction," Proceedings of the 4th International Conference on Cognitive Science and the 7th Australasian Society for Cognitive Science Joint Conference (ICCS/ASCS 2003), Vol. 1, 163-168, 2003.

[17] E. Wasserman, W. Dorner and S. Kao, "Contribution of specific cell information to judgments of interevent contingency," Journal of Experimental Psychology: Learning, Memory and Cognition 16(3), 509-521, 1990.

[18] J. Anderson and C. Sheu, "Causal inferences as perceptual judgements," Meomory and Cognition, 23, 510-524, 1995.

[19] M. Hattori and M. Oaksford, "Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis," Cognitive Science, 31(5), 765-814, 2007.

[20] H. Jenkins and W. Ward, "Judgment of contingency between responses and outcomes," Psychological Monographs: General and Applied, 79(1), 1-17, 1965.

[21] P. Cheng and L. Novick, "Covariation in natural causal induction," Psychological Review 99(2), 365-382, 1992.

[22] G. Gigerenzer and U. Hoffrage, "How to improve Bayesian reasoning without instruction: Frequency formats.," Psychol Rev 102: 684-704., 1995.

[23] D. Knill and A. Pouget, "The Bayesian brain: the role of uncertainty in neural coding and computation. Trends in Neurosciences 27: 712-719. Making.," London: Psychology Press., 2004.

[24] Manktelow, Thinking and Reasoning: An Introduction to the Psychology of Reason, Judgment and Decision, Psychology Press, 2012.

[25] E. Bonawits, T. Van Schijndel, D. Friel and L. Schulz, “Children balance theories and evidence in exploration, explanation and learning,” Cognitive Psychology, 64, 215–234., 2012.

[26] N. Goodman, T. Ullman and J. Tenenbaum, “Learning a theory of causality,” Psychological Review, 118, 110–119, 2011.

[27] A. Gopnik and H. Wellman, “Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory,” Psychological Bulletin, 138(6), 1085., 2012.

[28] T. Griffiths, E. Vul and A. Sanborn, “Bridging levels of analysis for probabilistic models of cognition,” Current Directions in Psychological Science, 21, 263–268, 2012.

[29] E. Pellicano and D. Burr, "When the world becomes 'too real': a Bayesian explanation of autistic perception," Trends in Cognitive Sciences, vol.16, No.10, pp504-510, 2012.

[30] J. Bowers and C. Davis, "Bayesian just-so stories in psychology and neuroscience," Psychological Bulletin, Vol. 138, No. 3, 389–414, DOI: 10.1037/a0026450, 2012.

[31] M. Jones and B. Love, "Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition," BEHAVIORAL AND BRAIN SCIENCES (2011) 34, 169–231, doi:10.1017/S0140525X10003134, 2011.

[32] M. Lee and B. Sarnecka, "Number-knower levels in young children: Insights from Bayesian modeling," Cognition 120 (2011) 391–402, doi:10.1016/j.cognition.2010.10.003, 2011.

[33] A. Perfors, J. Tenenbaum, T. Griffiths and F. Xu, "A tutorial introduction to Bayesian models of cognitive development," Cognition 120 (2011) 302–321, doi:10.1016/j.cognition.2010.11.015, 2011.

[34] C. Dawson and L. Gerken, "When global structure "Explains Away" local grammar: A Bayesian account of rule-induction in tone sequences," Cognition 120 (2011) 350–359, doi:10.1016/j.cognition.2010.11.016, 2011.

[35] C. Gallistel, S. Fairhurst and P. Balsam, "The learning curve: Implications of a quantitative analysis," Proceedings of the National Academy of Science, 101, 13124–13131, 2004.

[36] C. Gallistel, "The importance of proving the null," Psychological Review, 116, 439–453, 2009.

[37] S. Baker, A. Leslie, C. Gallistel and B. Hood, "Bayesian change-point analysis reveals developmental change in a classic theory of mind task," Cognitive Psychology 91 124–149, http://dx.doi.org/10.1016/j.cogpsych.2016.08.001, 2016.

[38] S. Papachristos and C. Gallistel, "Autoshaped head poking in the mouse: A quantitative analysis of the learning curve," Journal of the Experimental Analysis of Behavior, 85, 293–308, 2006.

[39] J. B. Brayanov and M. A. Smith, "Bayesian and "Anti-Bayesian" Biases in Sensory Integration for Action and Perception in the Size–Weight Illusion," Journal of Neurophysiology Published 1 March 2010 Vol. 103 no. 3, 1518-1531 DOI: 10.1152/jn.00814.2009 , 2010.

[40] X.-X. Wei and A. A. Stocker, "A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts," Nature Neuroscience 18, 1509–1517 (2015), doi:10.1038/nn.4105, 2015.

[41] F. Arecchi, "Phenomenology of Consciousness: from Apprehension to Judgment," Nonlinear Dynamics, Psychology and Life Sciences, 15-3, 359-375, 2011.

[42] P.-Y. Gunji, K. Sonoda and V. Basios, "Quantum cognition based on an ambiguous representation derived from a rough set approximation," Biosystems, 141, 55-66, 2016.

[43] Y. Gunji, M. Minoura, K. Kojima and Y. Horry, "Free will in Bayesian and inverse Bayesian inference-driven endo-consciousness," Progress in Biophysics and Molecular Biology, DOI: 10.1016/j.pbiomolbio.2017.06.018, 2017.

[44] O. E. Rossler, "Ultraperspective and endophysics," Biosystems Volume 38, Issues 2–3, 1996, Pages 211-219, https://doi.org/10.1016/0303-2647(95)01593-0, 1996.

[45] K. Matsuno, "From quantum measurement to biology via retrocausality," Progress in Biophysics and Molecular Biology, https://doi.org/10.1016/j.pbiomolbio.2017.06.012, 2017.

[46] Y. P. Gunji, T. Haruna and K. Sawa, "Principles of biological organization: Local–global negotiation based on "material cause"," Physica D: Nonlinear Phenomena Volume 219, Issue 2, 15 July 2006, Pages 152-167, https://doi.org/10.1016/j.physd.2006.06.001, 2006.

[47] K. Sasai and Y. P. Gunji, "Heterarchy in biological systems: A logic-based dynamical model of abstract biological network derived from time-state-scale re-entrant form," Biosystems Volume 92, Issue 2, May 2008, Pages 182-188, https://doi.org/10.1016/j.biosystems.2008.02.004, 2008.

[48] K. Sawa and Y. P. Gunji, "An emergence of formal logic induced by an internal agent," Biosystems Volume 124, October 2014, Pages 39-45, https://doi.org/10.1016/j.biosystems.2014.08.005, 2014.

[49] K. Sawa and A. U. Igamberdiev, "The Double Homunculus model of self-reflective systems," BioSystems 144 (2016) 1–7, http://dx.doi.org/10.1016/j.biosystems.2016.03.007, 2016.

[50] K. Sawa and A. U. Igamberdiev, "The dynamics of interaction of reflexive subjects operating with the two-valued versus many-valued logic," Progress in Biophysics and Molecular Biology. http://dx.doi.org/10.1016/j.pbiomolbio.2017.08.007, 2017.

[51] I. McAllister and D. T. Studlar, "Bandwagon, Underdog, or Projection? Opinion Polls and Electoral Choice in Britain, 1979-1987," The Journal of Politics Volume 53, Number 3 | Aug., 1991 , http://www.journals.uchicago.edu/doi/abs/10.2307/2131577, 1991.

[52] R. Macdonald and M. P. Taylor, "Exchange Rate Economics A Survey," IMF Economic Review March 1992, Volume 39, Issue 1, pp 1–57, https://doi.org/10.2307/3867200, 1992.

[53] L. Rosenkopf and E. Abrahamson, "Modeling Reputational and Informational Influences in Threshold Models of Bandwagon Innovation Diffusion," Computational & Mathematical Organization Theory December 1999, Volume 5, Issue 4, pp 361–384, https://doi.org/10.1023/A:1009620618662 , 1999.

[54] E. S. Asch, "Studies of Independence and Conformity: I.A Minority of One Against a Unanimous Majority," Psychological Monographs: General and Applied, 70(9), 1-70. http://dx.doi.org/10.1037/h0093718 , 1956.

[55] A. H. Simon, "Bandwagon and Underdog Effects and the Possibility of Election Predictions," Public Opinion Quarterly, Volume 18, Issue 3, 1 January 1954, Pages 245–253, https://doi.org/10.1086/266513, 1954.

[56] W. D. Fleitas, "Bandwagon and Underdog Effects in Minimal-Information Elections," American Political Science Review Volume 65, Issue 2 June 1971 , pp. 434-438 ,

https://doi.org/10.2307/1954459, 1971.

[57] C. Marsh, "Back on the Bandwagon The Effect of Opinion Polls on Public Opinion," British Journal of Political Science Volume 15, Issue 1 January 1985 , pp. 51-74 , https://doi.org/10.1017/S0007123400004063, 1985.

[58] E. F. Klor and E. Winter, "The welfare effects of public opinion polls," . International, Journal of Game Theory, 35 (3), 379-394. https://link.springer.com/article/10.1007%2Fs00182-006-0050-5?LI=true, 2007.

[59] J. Grosser and S. Arthur, "Public opinion polls, voter turnout, and welfare: an experimental study," American Journal of Political Science, 54 (3), 700-717. http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2010.00455.x/full, 2010.

[60] M. Agranov, J. K. Goeree, J. Romero and L. Yariv, "What makes voters turn out: The effects of polls and beliefs," Working Paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2029598, 2012.

[61] J. Duffy and M. Tavits, "Beliefs and voting decisions: A test of the pivotal voter model," American Journal of Political Science, 52 (3), 603-618. http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2008.00332.x/full, 2008.

[62] G. W. Cox, F. M. Rosenbluth and M. F. Thies, "Mobilization, Social Networks, and Turnout: Evidence from Japan," World Politics Volume 50, Issue 3 April 1998 , pp. 447-474, https://doi.org/10.1017/S0043887100012879, 1998.

[63] R. Nurkse, International Currency Experience: Lessons of the Interwar Period, Geneva: League of Nations, 1945.

[64] H. Allen and M. P. Taylor, "Charts, noise and fundamentals in the London foreign exchange market," The Economic Journal", Vol. 100, No. 400, Conference Papers (1990), pp. 49-59, http://www.jstor.org/stable/2234183 , 1990.

[65] H. Siddiqi, "Stock Price Manipulation: The Role of Intermediaries," International Journal of Financial Studies. 2017, 5(4), 24; doi:10.3390/ijfs5040024, 2007a.

[66] H. Siddiqi, "Stock Price Manipulation: The Role of Intermediaries," CMER WORKING PAPER No. 07-58, http://saber.eaber.org/node/22280, 2007b.

[67] E. Abrahamson, "Managerial Fads and Fashions: The Diffusion and Rejection of Innovations," Academy of Management Review, 16, 586–612. http://amr.aom.org/content/16/3/586.full.pdf+html, 1991.

[68] E. Abrahamson and L. Rosenkopf , "Institutional and Competitive Bandwagons: Using Mathematical Modeling as a Tool to Explore Innovation Diffusion," Academy of Management Review, 18, 487–517.. doi: 10.5465/AMR.1993.9309035148, 1993.

[69] S. Moscovici and M. Zavalloni, "The group as a polarizer of attitudes," Journal of Personality and Social Psychology, 12, 125-135, 1969.

[70] G. Mugny and S. Papastamou, "When rigidity does not fail: Individualization and psychologization as resistances to the diffusion of minority innovations.," European

Journal of Social Psychology, 10(1), 43-61. , 1980.

[71] C. J. Nemeth, "Differential contributions of majority and minority influence," Psychological Review, 93, 23-32. , 1986.

[72] M. R. Trost, A. Maass and D. T. Kenrick, "Minority influence: Personal relevance biases cognitive processes and reverses private acceptance," Journal of Experimental Social Psychology, 28,234-254., 1992.

[73] H. W. Chappell, R. R. McGregor and T. A. Vermilyea, "Learning and Bandwagons in Monetary Policy Committees," Southern Economic Association 2010 Sessions for Rob Roy McGregor Session 14F Monetary Policy Committee Decisions, http://professorchappell.com/Papers/Chappell_etal_Learning_Bandwagons_10-09-2009.pdf, 2010.

[74] 堀井洋一, 吉成愛, 中本百合菜 , 郡司ペギオ幸夫, “逆ベイズ推定を用いた意思決定のモデル化 ,” 計測自動制御学会論文誌 Vol.54 No.1 31-38, https://doi.org/10.9746/sicetr.54.31, 2018.

[75] Y. Horry, "Financial information description language and visualization and analysis tools," Computer Languages, Systems & Structures, 50 (2017) 31–52, http://dx.doi.org/10.1016/j.cl.2017.05.005, 2017.

[76] M. Tainaka, "Treasure in the Depth of Ocean," Hitachi Review, Vol.63, No.3, pp.24-29, 2014.

[77] R. Panko, "What we know about spreadsheet errors," Journal of End User Computing, vol. 10, pp. 15–21, 1998.

[78] E. McLean, L. Kappelman and J. Thompson, "Converging End-User and Corporate Computing," Communications of the ACM (36:12) December 1993, pp. 79-92, 1993.

[79] B. Boehm,, C. Abts,, A. Brown,, S. Chulani,, K. Bradford,, E. Horowitz,, R. Madachy,, D. Reifer and B. Steece,, Software Cost Estimation with COCOMO, Prentice-Hall International, Upper Saddle River, NJ, 2000.

[80] R. Panko, "Spreadsheet Errors: What We Know. What We Think We Can Do.," Proceedings of the Spreadsheet Risk Symposium, European Spreadsheet Risks Interest Group (EuSpRIG), 2000.

[81] G. Gable, C. Yap and M. Eng, "Spreadsheet Investment, Criticality, and Control,," Proceedings of the Twenty-Fourth Hawaii International Conference on System Sciences, Vol. III, Los Alomitos, CA: IEEE Computer Society Press, 1991, pp. 153-162, 1991.

[82] Y. Chan and V. Storey, "The Use of Spreadsheets in Organizations: Determinants and Consequences," Information & Management, (31:3) December 1996, pp. 119-134, 1996.

[83] P. Cragg and M. King, "Spreadsheet Modelling Abuse: An Opportunity for OR?," Journal of the Operational Research Society (44:8) August 1993, pp. 743-752, 1993.

[84] B. Floyd, J. Walls and K. Marr, "Managing Spreadsheet Model Development," Journal of Systems Management 46:1) May/June 1995, pp. 38-43, 68, 1995.

[85] M. Hall, "A Risk and Control Oriented Study of the Practices of Spreadsheet Application Developers," Proceedings of the Twenty-Ninth Hawaii International Conference on Systems Sciences, Vol. II, Kihei, Hawaii, Los Alamitos, CA: IEEE Computer Society Press, January 1996, pp. 364-373, 1996.

[86] T. Teo and M. Tan, "Quantitative and Qualitative Errors in Spreadsheet Development," In 30th Hawaii Int. Conf. on System Sciences, pages 25-38, 1997.

[87] K. Rajalingham, D. Chadwick, B. Knight and D. Edwards, "Quality Control in Spreadsheets: A Software Engineering-Based Approach to Spreadsheet Development," 33rd Hawaii International Conference on System Sciences, pages 1–9, 2000.

[88] M. Tukiainen, "Uncovering Effects of Programming Paradigms: Errors in Two Spreadsheet Systems," 12th Workshop of the Psychology of Programming Interest Group (PPIG), pages 247–266, 2000.

[89] K. Rajalingham, D. Chadwick and B. Knight, "Classification of Spreadsheet Errors," European Spreadsheet Risks Interest Group (EuSpRIG), 2001.

[90] R. Panko and S. Aurigemma, "Revising the Panko–Halverson taxonomy of spreadsheet errors," Decision Support Systems Volume 49, Issue 2, May 2010, Pages 235–244, 2010.

[91] R. Abraham and M. Erwig, "Header and unit inference for spreadsheets through spatial analyses," in Visual Languages and Human Centric Computing, 2004 IEEE Symposium, pp. 165-172, 2004.

[92] S. Roy, F. Hermans, E. Aivaloglou, J. Winter and A. van Deursen, "Evaluating Automatic Spreadsheet Metadata Extraction on a Large Set of Responses from MOOC Participants," IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER) pp.135-145, 2016.

[93] *Wikipedia Lotus Improv, https://en.wikipedia.org/wiki/Lotus_Improv.*

[94] M. Erwig and M. Burnett, "Adding Apples and Oranges," In 4th Int. Symp. on Practical Aspects of Declarative Languages LNCS 2257, pages 173–191, 2002.

[95] M. Burnett and M. Erwig, "Visually Customizing Inference Rules About Apples and Oranges," In 2nd IEEE Int.Symp. on Human-Centric Computing Languages and Environments pages 140–148, 2002.

[96] Y. Ahmad, T. Antoniu, S. Goldwater and S. Krishnamurthi, "A Type System for Statically Detecting Spreadsheet Errors," In 8th IEEE Int. Conf. on Automated Software Engineering, pp.174–183, 2003.

[97] M. Burnett, C. Cook, J. Summet, G. Rothermel and C. Wallace, "End-user Software Engineering with Assertions," 25th Int. Conference on Software Engineering, 2003.

[98] T. Antoniu, P. Steckler, S. Krishnamurthi, E. Neuwirth and M. Felleisen, "Validating the Unit Correctness of Spreadsheet Programs," In Int. Conf. on Software Engineering, 2004.

[99] R. Abraham and M. Erwig, "Ucheck: A spreadsheet type checker for end users," Journal of Visual Languages & Computing, vol. 18, no. 1, pp. 71–95, 2007.

[100] M. Erwig and M. Burnett, Adding apples and oranges, in Practical Aspects of Declarative Languages. Springer, 2002, pp. 173–191, 2002.

[101] F. Hermans, M. Pinzger and A. van Deursen, Automatically extracting class diagrams from spreadsheets, in ECOOP 2010–Object-Oriented Programming. Springer pp. 52–75, 2010.

[102] J. Cunha, J. Saraiva and J. Visser, "From spreadsheets to relational databases and back," in Proceedings of the 2009 ACM SIGPLAN Workshop on Partial Evaluation and Program Manipulation, ser. PEPM '09. New York, NY, USA: ACM, 2009, pp. 179–188, 2009.

[103] M. Vos, W. Hage, J. Ros and G. Schreiber, "Reconstructing Semantics of Scientific Models : a Case Study," In: Proceedings of the OEDW workshop on ontology engineering in a data driven world, EKAW pp.79-86, 2012.

[104] S. Chatvichienchai, Spreadsheet metadata extraction: A layout-based approach, in Database and Expert Systems Applications. Springer pp. 147–160, 2012.

[105] S. Chatvichienchai, "Automatic metadata extraction and classification of spreadsheet documents based on layout similarity," in Advanced Information Management and Service (ICIPM), 7th International Conference on, Nov 2011, pp. 38–43, 2011.

[106] Z. Chen and M. Cafarella, "Automatic web spreadsheet data extraction," in Proceedings of the 3rd International Workshop on Semantic Search over the Web. ACM, 2013, p. 1, 2013.

[107] T. Isakowitz, S. Schocken and H. Lucas, Jr, "Toward a Logical/Physical Theory of Spreadsheet Modelling," ACM Transactions on Information Systems, 13(1):1–37, 1995.

[108] A. Yoder and D. Cohn, "Real Spreadsheets for Real Programmers," Int. Conference on Computer Languages, pages 20–30, 1994.

[109] B. Ronen, M. Palley and H. Lucas, Jr, "Spreadsheet Analysis and Design," Communications of the ACM, 32(1):84–93, 1989.

[110] A. Blackwell and T. Green, "Notational Systems -The Cognitive Dimensions of Notations Framework," HCI Models, Theories, and Frameworks: Toward and Interdisciplinary Science, Morgan Kaufmann Publishers, pages 103–133, 2003.

[111] G. Rothermel, M. Burnett, L. Li, C. Dupuis and A. Sheretov, "A Methodology for Testing Spreadsheets," ACM Transactions on Software Engineering and Methodology, vol.10, pages 110–147, 2001.

[112] F. Hermans, M. Pinzger and A. Van Deursen, "Supporting professional spreadsheet users by generating leveled dataflow diagrams," in Proceedings of the 33rd International Conference on Software Engineering ACM, 2011, pp. 451–460, 2011.

[113] M. Mernik, J. Heering and A. Sloane, "When and how to develop domain-specific languages," ACM Computing Surveys, 37(4), 316-344, 2005.

[114] T. Kosar, S. Bohra and M. Mernik, "Domain-Specific Languages: A Systematic Mapping Study," Information and Software Technology, Volume 71, March 2016, Pages 77–91, 2016.

[115] T. Kosar, M. Mernik and J. Carver, “Program comprehension of domain-specific and general-purpose languages: Comparison using a family of experiments,” Empirical Software Engineering 17 (3) (2012) 276-304, 2012.

[116] S. Melia, C. Cachero, J. Hermida and E. Aparicio, “Comparison of a textual versus a graphical notation for the maintainability of MDE domain models: an empirical pilot study,” Software Quality Journal, 24(3), pp.709–735, 2016.

[117] F. Häser, M. Felderer and R. Breu, “Is Business Domain Language Support Beneficial for Creating Test Case Specifications: A Controlled Experiment,” Information and Software Technology, Volume 79, November 2016, pp. 52–62, 2016.

[118] A. Johanson and W. Hasselbring, “Effectiveness and efficiency of a domain-specific language for high-performance marine ecosystem simulation: a controlled experiment,” Empirical Software Engineering, 22 November 2016, pp. 1-31, DOI: 10.1007/s10664-016-9483-z, 2016.

[119] R. Kieburtz, L. McKinney, J. Bell, J. Hook, A. Kotov, J. Lewis, D. Oliva, T. Sheard, I. Smith and L. Walton, “A software engineering experiment in software component generation.,” Proceedings of the 18th international conference on software engineering (ICSE’96), pp 542–552, 1996.

[120] Y. Horry, A. Yoshinari, Y. Nakamoto and Y. P. Gunji, "Modeling of decision-making process for moving straight using inverse Bayesian inference," Biosystems Volume 163, January 2018, Pages 70–81, https://doi.org/10.1016/j.biosystems.2017.12.006, 2017.

[121] E. A. Liberman, S. V. Minina, D. A. Moshkov, I. M. Santalova, I. A. Chistopolskiy and N. E. Shklovski-Kordi, "Experimental testing of the role of cytoskeleton in the solution by neurons of problems facing the brain," Biochemistry (Moscow), 2008, Vol. 73, No. 4, pp. 479-482. DOI: 10.1134/S0006297908040147, 2008.

[122] A. U. Igamberdiev and N. E. Shklovskiy-Kordi, "The quantum basis of spatiotemporality in perception and consciousness," Progress in Biophysics and Molecular Biology 130 (2017) 15-25, http://dx.doi.org/10.1016/j.pbiomolbio.2017.02.008, 2017.

[123] A. Thomas and B. Oommen, Optimal “Anti-Bayesian” Parametric Pattern Classification Using Order Statistics Criteria, In: Alvarez L., Mejail M., Gomez L., Jacobo J. (eds) Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2012. Lecture Notes in Computer Science, vol 7441. Springer, Berlin, Heidelberg, 2012a.

[124] A. Thomas and B. Oommen, Optimal “Anti-Bayesian” Parametric Pattern Classification for the Exponential Family Using Order Statistics Criteria, In: Campilho A., Kamel M. (eds) Image Analysis and Recognition. ICIAR 2012. Lecture Notes in Computer Science, vol 7324. Springer, Berlin, Heidelberg, 2012b.

[125] A. Thomas and J. B. Oommen, "The Fundamental Theory of Optimal “Anti-Bayesian” Parametric Pattern Classification Using Order Statistics Criteria," Pattern Recognition Volume 46, Issue 1, January 2013, Pages 376-388, 2013.

[126] "HuffPost Pollster," 2016. [Online]. Available: http://elections.huffingtonpost.com/pollster/2016-general-election-trump-vs-clinton.csv.

[127] “Dow-Jones Average,” 2017. [Online]. Available: https://jp.investing.com/indices/us-30-historical-data.

[128] “Exchange rate,” 2017. [Online]. Available: https://www.mizuhobank.co.jp/rate/market/historical.html.

[129] “Oil price,” 2017. [Online]. Available: https://jp.investing.com/commodities/crude-oil-historical-data.

[130] L. Li, A. Scaglione , A. Swami and Q. Zhao, "Phase Transition in Opinion Diffusion in Social Networks," Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, DOI: 10.1109/ICASSP.2012.6288564, 2012.

Figure and Table Captions

Publications

Journals

1. 〇**<u>堀井洋一</u>**、吉成愛、中本百合菜、郡司ペギオ幸夫 (2018)、"逆ベイズ推定を用いた意思決定のモデル化"、計測自動制御学会論文誌 Volume **54** No.1 31-38, https://doi.org/10.9746/sicetr.54.31
2. 〇**<u>Youichi Horry</u>**, Ai Yoshinari, Yurina Nakamoto, Yukio-Pegio Gunji (2018), "Modeling of decision-making process for moving straight using inverse Bayesian inference", BioSystems Volume **163** 70-81 (2018/1), https://doi.org/10.1016/j.biosystems.2017.12.006
3. 〇**<u>Youichi Horry</u>** (2017), "Financial Information Description Language and Visualization / Analysis Tools", Computer Languages, Systems & Structures **50** (2017) 31-52. http://dx.doi.org/10.1016/j.cl.2017.05.005

4. Yukio-Pegio Gunji, Mai Minoura, Kei Kojima, **<u>Yoichi Horry</u>** (2017), "Free will in Bayesian and inverse Bayesian inference-driven endo-consciousness", Progress in Biophysics and Molecular Biology Volume **131** 312-324 (2017/12), http://dx.doi.org/10.1016/j.pbiomolbio.2017.06.018
5. 三宅吉宣、市毛由希子、林政一郎、黒川秀昭、**<u>堀井洋一</u>**、影山晃冶 (2013)、"大型海水淡水化事業への日立グループの取り組み", 日立評論 Volume **95** No.08 562-563 http://www.hitachihyoron.com/jp/pdf/2013/08/2013_08_10.pdf
6. 三木和穂、柳邦宏、古谷純、**<u>堀井洋一</u>** (2007), "放送と通信の融合・連携時代への取り組み—新しいライフスタイルの実現に向けて—"、日立評論、Volume **89** No.06 464-465 http://www.hitachihyoron.com/jp/pdf/2007/06/2007_06_02.pdf
7. Hiroko Sukeda, **<u>Youichi Horry</u>**, Yukinobu Maruyama, Takeshi Hoshino (2007), "Human Interfaces Supporting New Encounters with Digital Content", Hitachi Review, Volume **56** 19-23 http://www.hitachi.com/rev/pdf/2007/r2007_01_105.pdf
8. 助田浩子、**<u>堀井洋一</u>**、丸山幸伸、星野剛史 (2006), "コンテンツとの新しい出会いを誘発する ヒューマンインタフェース", 日立評論 Volume **88** No.3 259-262 http://www.hitachihyoron.com/jp/pdf/2006/03/2006_03_06.pdf
9. Rieko Otsuka, Takeshi Hoshino, **<u>Youichi Horry</u>** (2006), "Transpost: A Novel Approach to the Display and Transmission of 360 Degrees-Viewable 3D Solid Images", IEEE Transactions on Visualization and Computer Graphics Volume **12** Issue 2 178-185 (2006/1/23), http://dx.doi.org/10.1109/TVCG.2006.38
10. 助田浩子、大関一博、**<u>堀井洋一</u>** (2006), "Anonymous CRM ----顧客の匿名性を考慮した顧客情報分析システム"、情報処理学会論文誌 **47**(3) 658-666 (2006/3/15), https://ipsj.ixsq.nii.ac.jp/ej/?action=repository_action_common_download&item_id=10350&item_no=1&attribute_id=1&file_no=1

11. **Youichi Horry**, Isshu Nakajima, Takeshi Hoshino, Yukinobu Maruyama (2003), "A Passive-style buttonless mobile terminal", IEEE Transactions on Consumer Electronics, Volume **49**, No. 3 530-535, (2003/8) http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1218778
12. Samuel Boivin, Andre Gagalowicz, **Youichi Horry** (2000), "Morphological Modeling and Deformation of 3D objects", in International Journal Machine Graphics and Vision, **9**(1/2), 263-280, Warsaw, Poland. http://www.dgp.toronto.edu/~boivin/pubs/morpho.pdf
13. **堀井洋一**、新井清志、野村洋之、安生健一 (1997), "Tour Into the Picture －1枚の絵からのアニメーションー", 画像ラボ **8** 巻 12 号 1-5 (1997/12), http://jglobal.jst.go.jp/public/20090422/200902189204430672
14. **Youichi Horry**, Ken Anjyo, Kiyoshi Arai (1997), "Tour into the picture: using a spidery mesh interface to make animation from a single image". Proceedings SIGGRAPH '97 225-232 (Los Angels, California 1997/8/3-8/8), http://research.cs.queensu.ca/~jstewart/850/papers/p225-horry.pdf
15. 今野紀雄、**堀井洋一**、荒井寿勝、石堂正人、中浜清志 (1990)、"音高列としての音楽に関するパワースペクトル(特に 1/f ノイズ)とフラクタル次元について", 電子情報通信学会論文誌 D Volume **J73-D2** No. 4 495-500 (1990/04/20), http://search.ieice.org/bin/author.php?a_name=%C0%D0%C6%B2%A1%A1%C0%B5%BF%CD&category=D&year=1990&lang=J&abst=
16. 今野紀雄、**堀井洋一**、中浜清志 (1989), "作曲アルゴリズム"Profile"とのその周辺 Music compositional algorithm "Profile" and its circumference", Computer Today 6 巻 **6** 号 37-41 http://jglobal.jst.go.jp/public/20090422/200902030277506775
17. **Youichi Horii**, Naoyuki Fujii; Hiroshi Takeda (1990), "Hardness analysis of metallic particles in ordinary chondrites", Proceedings of the NIPR Symposium on Antarctic Meteorites **3** 254-263 (1990/10), http://ci.nii.ac.jp/els/110000029914.pdf?id=ART0000355623&type=pdf&lang=en&host=cinii&order_no=&ppv_type=0&lang_sw=&no=1331621045&cp=
18. **堀井洋一**、下岡広志、浅山勇介、森永速男、庫本 正 (1990), "秋芳洞における石灰華段丘(百枚皿)の形状解析とその成因", 日本洞窟学会第 **15** 巻 34-41(1990/10/31)

International conference

1. Satomi Tsuji, Koji Ara, Nobuo Sato, Yoshihiro Wakisaka, Kazuo Yano, Norio Ohkubo, Rieko Otsuka, Miki Hayakawa, Norihiko Moriwaki, **Youichi Horry** (2009), "Visualization of knowledge-creation process using face-to-face communication data", 2009 Sixth International Conference on Networked Sensing Systems (INSS), http://dx.doi.org/10.1109/INSS.2009.5409922
2. Koji Ara, Nobuo Sato, Satomi Tsuji, Yoshihiro Wakisaka, Norio Ohkubo, **Youichi**

Horry, Norihiko Moriwaki, Kazuo Yano, Miki Hayakawa (2009), "Predicting flow state in daily work through continuous sensing of motion rhythm", Networked Sensing Systems (INSS), 2009 Sixth International Conference on, 1-6 (2009/6/17), http://dl.acm.org/citation.cfm?id=1802340.1802377&coll=DL&dl=GUIDE

3. Yoshihiro Wakisaka, Norio Ohkubo, Koji Ara, Nobuo Sato, Miki Hayakawa, Satomi Tsuji, **Youichi Horry**, Kazuo Yano, Norihiro Moriwaki (2009), "Beam-scan sensor node: reliable sensing of human interactions in organization", 2009 Sixth International Conference on Networked Sensing Systems, http://dl.acm.org/citation.cfm?id=1802340.1802352&coll=DL&dl=GUIDE
4. Jun-ichiro. Watanabe, **Youichi Horry** (2009), "Memorization and Information-Retrieval Behaviors". In: Gross T. et al. (eds) Human-Computer Interaction – INTERACT 2009. INTERACT 2009. Lecture Notes in Computer Science, vol 5726. Springer, Berlin, Heidelberg https://link.springer.com/chapter/10.1007/978-3-642-03655-2_51
5. Jun-ichiro Watanabe, Arito Mochizuki, **Youichi Horry** (2008), "Bookisheet: bendable device for browsing content using the metaphor of leafing through the pages", Conference: UbiComp 2008: Ubiquitous Computing, 10th International Conference, UbiComp 2008, (Seoul, Korea 2008/9/21 – 9/24), http://dx.doi.org/10.1145/1409635.1409684
6. Hiroko Sukeda, Arito Mochizuki, Isshu Nakajima, **Youichi Horry** (2008), "Cocktail Interface: Server-based User Interface for Seamless Access to Various Kinds of Digital Content", Proceedings of the International Symposium on Consumer Electronics, ISCE (2008), https://doi.org/10.1109/ISCE.2008.4559563
7. Rieko Otsuka, Takeshi Hoshino, **Youichi Horry** (2007), "360 degrees-viewable Display of 3D Solid Images", ACM SIGGRAPH 2007 Poster, https://www.cct.lsu.edu/~fharhad/ganbatte/siggraph2007/CD1/content/posters/0102.pdf
8. Hiroko Sukeda, Yujin Tsukada, Yukinobu Maruyama, Isshu Nakajima, Takeshi Hoshino, Jun-ichiro Watanabe, **Youichi Horry** (2007), "Wrapping TV Interface for Watching Multiple Programs Efficiently", Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on 1-2
9. Rieko Otsuka, Takeshi Hoshino, **Youichi Horry** (2006), "Transpost: 360 deg-viewable three-dimensional display system", Proceedings of the IEEE, Volume 94 Issue 3, 629-635 (2006/3), http://dx.doi.org/10.1109/JPROC.2006.870700
10. Hiroko Sukeda, **Youichi Horry**, Yukinobu Maruyama, Takeshi Hoshino (2006), "Information-accessing furniture to our everyday lives more comfortable", 2006 Digest of Technical Papers International Conference on Consumer Electronics http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1598292
11. Tadaaki Ishikawa, **Youichi Horry**, Takeshi Hoshino (2005), "Touchless input device and gesture commands", Consumer Electronics, 2005. ICCE. 2005 Digest of Technical Papers. International Conference on 205-206 (2005/1/8), http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1429789&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D1429789

12. Rieko Otsuka. Takeshi Hoshino, **Youichi Horry** (2005), "Transpost: 360°-viewable Three-Dimensional Display, System", Proceedings of the IEEE (3-D Technologies for Imaging and Display) http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel5%2F5%2F33728%2F01605207.pdf%3Ftp%3D%26arnumber%3D1605207%26punumber%3D5&authDecision=-203
13. Rieko Otsuka. Takeshi Hoshino, **Youichi Horry** (2004), "Transpost: All-Arount Three-Dimensional Display System", SPIE Optical East 2004 Three-Dimensional TV, Video, and Display, Volume 5599 56-66. http://scholar.google.co.jp/citations?view_op=view_citation&hl=en&user=EBQOU2YAAAAJ&citation_for_view=EBQOU2YAAAAJ:Tyk-4Ss8FVUC
14. Hiroko Sukeda, Kazuhiro Ozeki, Chiaki Matsumoto, **Youichi Horry**, "Implementation of Instant CRM: Building One-to-One Relationship with Loyal Customers", INFORMS Marketing Science Conference 2004 25-25 FA10
15. Rieko Otsuka. Takeshi Hoshino, **Youichi Horry** (2004), "Transpost: all-around display system for 3D solid image", Proceedings of ACM Symposium on Virtual Reality Software and Technology 2004, http://dl.acm.org/citation.cfm?id=1077534.1077576&coll=DL&dl=GUIDE
16. Ken Anjyo, Yoshiyuki Momose, **Youichi Horry**, Michael Arias (2000), "Digial Animation in Japan", ACM SIGGRAPH 2000 Panel, http://www.siggraph.org/s2000/conference/panels/panels4.html
17. **Youichi Horry** (1998), "Crash Mapping", ACM SIGGRAPH SiggraphTV, http://www.siggraph.org/s98/conference/stv/proj.html
18. Ken Anjyo, **Youichi Horry** (1998), "Theory and Practice of "Tour Into the Picture" ACM SIGGRAPH 1998 Course, http://www.cc.gatech.edu/classes/AY2000/cs4480_spring/cases/tip-siggraph98/html/tip000.htm
19. **Youichi Horry**, Ken Anjyo, Kiyoshi Arai (1997), "Tour Into the Picture", ACM SIGGRAPH Electric Theatre, Video Review 110
20. **Youichi Horry** (1994), "Graphical User Interface for MIDI Signal Generation and Sound Synthesis", Proceedings of the International Computer Music Conference ICMC `94 276-279 (Aarhus 1994/9), http://quod.lib.umich.edu/i/icmc/bbp2372.1994.072/1
21. Hinoharu Matsumoto, **Youichi Horry**, Masako Kawamura, Akikazu Nakamura (1994), "Dawn Bird（あかつきの鳥）", Music From Japan, The Asia Society, (New York 1994/3/19)
22. Naoyuki Fujii, **Youichi Horii**, Hiroshi Takeda (1990), "Hardness and Shape Irregurarity of Metallic Particles in Antarctic Ordinary Chondrites", Abstracts of the Lunar and Planetary Science Conference, Volume 21, 395, http://adsabs.harvard.edu/full/1990LPI....21..395F
23. **Youichi Horii**, Naoyuki Fujii, Hiroshi Takeda (1989), "Mechanical and Morphological Characterization of Shock Effects in Antarctic Meteorites (2) Hardness & Grain Shape

Analysis of Metallic Particles, Antarctic Meteorites XIV". Papers presented to the 14th Symposium on Antarctic Meteorites, NIPR, 78-80 (Tokyo 1989/7/6 – 7/8) http://adsabs.harvard.edu/abs/1989anme...14...78H

24. Naoyuki Fujii, **Youichi Horii**, Hiroshi Takeda (1988), "Mechanical and Morphological Characterization of Shock Effects in Antarctic Meteorites (1)", Antarctic Meteorites XIII. Papers presented to the 13th Symposium on Antarctic Meteorites, NIPR 107 (Tokyo 1988 6/7 -6/9), http://adsabs.harvard.edu/abs/1988anme...13..107F
25. Naoyuki Fujii, **Youichi Horii**, Masamichi Miyamoto (1988), "Metallic Particles in Ureilites -- Preliminary Report from Grain Shape Analysis." Proceedings of ISAS Lunar and Planetary Symposium (21st) 114-119, https://gbank.gsj.jp/geolis/geolis_link/199100529/en

Domestic conference

1. ○堀井洋一 (2018)、"逆ベイズ推論が決める、勝ち馬に乗るか否か"、第 12 回内部観測研究会＋第 29 回計測自動制御学会 SI 部門共創システム部会研究会, (2018/3)
2. ○堀井洋一、吉成愛、中本百合菜、郡司ペギオ幸夫 (2017)、"直進歩行を阻む脳:逆ベイズ推定を用いた意思決定のモデル化"、第 11 回内部観測研究会＋第 28 回計測自動制御学会 SI 部門共創システム部会研究会 (2017/2/25) http://www.ypg.ias.sci.waseda.ac.jp/IMS2017/index.html
3. ○堀井洋一、吉成愛、中本百合菜、郡司ペギオ幸夫 (2016)、"逆ベイズ推定を用いた意思決定のモデル化"、第 17 回計測自動制御学会システムインテグレーション部門講演会 (札幌 2016/12/15):優秀講演賞受賞

4. 荒宏視、秋富知明、福間晋一、渡邊純一郎、佐藤信夫、森脇紀彦、矢野和男、**堀井洋一**、守屋俊夫 (2012), "センサ活用による行動変革 Using Sensing Technology for Changing Behavior"日立製作所 中央研究所 回路とシステムワークショップ
5. 大塚理恵子, 星野剛史, 堀井洋一 (2007), "小型で持ち運び可能な 360 度立体映像ディスプレイ技術" 映像情報メディア学会 立体映像技術研究会 情報ディスプレイ研究会, http://d.hatena.ne.jp/rittaikyo/mobile?sid=167f12f95372093d&date=20071017&section=1186645403
6. 石川忠明、**堀井洋一**、星野剛史 (2005), "タッチレス・インターフェース・デバイスとジェスチャーコマンド", 映像情報メディア学会技術報告 29(21), 13-16 (2005/3/11), http://ci.nii.ac.jp/naid/10015700276

7. 大塚理恵子、堀井洋一、星野剛史 (2005) “Transpost: 360 度立体映像ディスプレイシステム”、JAPAN AVS Consortium 第 11 回ビジュアリゼーションカンファレンス, http://kgt.cybernet.co.jp/avs_conso/event/vc11/summary/data/1-2.pdf
8. 石井隆昭、望月有人、星野剛史、堀井洋一 (2005), “IKESU:「旬」な音楽を聴くための収集型ミュージックプレーヤー”, WISS 2005 (2005/12/7 – 12/9), http://www.wiss.org/WISS2005/program.html
9. 大塚理恵子, 星野剛史, 堀井洋一 (2005), “入力をやわらかく支援する布キーボード”, WISS 2005 (2005/12/7 – 12/9), http://www.wiss.org/WISS2005/program.html
10. 大塚理恵子, 星野剛史, 堀井洋一 (2005) , “Transpost:360 度立体映像ディスプレイシステム”, 電気学会 電子・情報・システム部門
11. 土井美鈴, 堀井洋一 (2004), “プライバシーを考慮した監視カメラ映像配信システム”, インタラクション 2004, http://www.interaction-ipsj.org/archives/paper2004/pdf2004/B14.pdf
12. 大塚理恵子, 星野剛史, 堀井洋一, 柳本学, 丸山幸伸 (2004), “Transpost: 360 度どこからでも見ることができる立体映像ディスプレイ”, WISS 2004 (2004/12/1 – 12/3), http://www.wiss.org/WISS2004/program.html
13. 助田浩子、大関一博、松本千晶、堀井洋一 (2003), “Instant CRM on Mobile — 携帯電話を 用いた顧客情報分析システム”, マルチメディア分散、協調とモバイル (dicomo 2003) シンポジウム論文集 681-684 (2003/6/6), http://css.jaist.ac.jp/kfdb/reports_pub_if_no_claim_we_can_delete_it/tpdf/ipsj2003dicomo/170.pdf
14. 星野剛史, 堀井洋一, 丸山幸伸, 片山淳詞, 柴田吉隆, 吉丸卓志 (2001), “Air Real: ホームネットワークのユーザーインタフェース”, WISS 2001 (2001/12/5 – 12/7), http://www.wiss.org/WISS2001/program.html
15. 丸山幸伸, 堀井洋一, 星野剛史, 中島一州 (2001), “Waterscape:ジェスチャー入力による受動型モバイル情報端末”, WISS 2001 (2001/12/5 – 12/7), http://www.wiss.org/WISS2001/program.html
16. 安生健一、堀井洋一、新井清志 (1997), “Tour Into the Picture: 1枚の画像に基づくアニメーション生成技術”, Visual Computing 97 pp.31-34
17. 新井清志、堀井洋一 (1996), “二次元ドローデータによる三次元多関節構造体のモデリング”, 情報処理学会研究報告グラフィクスと CAD（CG） 1996(18(1995-CG-079)), 15-22 (1996/2/23), http://ci.nii.ac.jp/naid/110002781007
18. 堀井洋一、新井清志 (1996), “任意の視点移動を実現する対話的画像再合成”, 情報処理学会研究報告グラフィクスと CAD（CG） 1996(125(1996-CG-083)), 37-42 (1996/12/19), http://ci.nii.ac.jp/naid/110003274599
19. 堀井洋一 (1995)、”コンピュータ音楽の研究動向” Compaq Media Salon 1995/10
20. 松本日之春、堀井洋一、古橋潤一、加納麻衣子 (1993), “凍れる鳥”, 電楽 II 銀座十字屋ホール (1993/2/11)
21. 堀井洋一 (1992), “セル・オートマトンを用いた音楽の認知過程モデル”, 計測制御学会第19回パターン計測部会研究会 (川崎 NKK 京浜ビル 1992/6/4)
22. 松本日之春、堀井洋一、古橋潤一、高橋律也 (1992)、”風の鳥 L'oiseau du vent”,

愛知県芸術劇場コンサートホール (1992/12/16)
23. 松本日之春、**堀井洋一**、草間路代 (1992), “古の鳥 L'oiseau d'autrefois”, 日本電子音楽協会 第1回定期演奏会 (1992/11/3)
24. 松本日之春、**堀井洋一** (1991), “ミディア・見えない鳥”, 音楽情報科学研究会(共催 日本現代音楽協会) 第 35 回例会:東京現代音楽祭 電楽 コンピュータ音楽の現在・未来展第 1 部 コンピュータ ミュージック エキシビジョン——先端テクノロジーとアートの融合・21 世紀における音楽創造電楽 東京神楽坂 音楽の友ホール (1991/12/8), http://nagasm.org/ASL/jmacs/index.html
25. **堀井洋一** (1990), “表象主義以降の認知モデル——聴音の錯覚を例として”, 音楽情報科学研究会 夏のシンポジウム (立教大学他 1990/8/31-9/2), http://nagasm.org/ASL/jmacs/index.html
26. **堀井洋一** (1989), “音楽と時間”, 第1回観察倶楽部 東大医学部
27. **堀井洋一** (1989), “音楽に意味を与える形式”, 数理物理研究会 室蘭工業大学
28. **堀井洋一** (1989), “Copy Feedback”, 数理物理研究会 室蘭工業大学
29. **堀井洋一** (1988), “両眼視差無しの実体視”, 第2回観察倶楽部 琵琶湖さざなみ荘、
30. 藤井直之、**堀井洋一** (1989), “隕石から見た微惑星の衝突過程”, 文部省科学研究費補助金重点領域(I)「原始太陽系と惑星の起源」 307-317
31. 藤井直之、**堀井洋一**、金丸剛、三軒一義 (1989), “オリビン中の転位組織と衝撃の影響”, 月惑星シンポジウム 宇宙科学研究所
32. **堀井洋一**、藤井直之、武田弘 (1989), “隕石における衝撃の影響-2(演旨)”, 地震学会講演予稿集, 1989, 1, 280-280, https://gbank.gsj.jp/geolis/geolis_link/198901132/ja
33. **堀井洋一**、藤井直之、武田弘 (1988), “隕石における衝撃の影響(I)”, 地震学会秋期大会予稿集 152 B66
34. 中野司、藤井直之、**堀井洋一** (1988), “GRAIN”と“EDGE”--粒子画象解析のためのソフトウェア(演旨)GRAIN and EDGE Softwares for Grain Image Analysis”, 情報地質研究会 1988 巻 (1988) 13 号 93-117, https://doi.org/10.6010/geoinformatics1975.1988.13_93
35. **堀井洋一**、藤井直之、武田弘 (1988), “コンドライトにおける衝撃の影響(演旨)”, 地震学会講演予稿集 1988 巻 2 号 152-152, https://gbank.gsj.jp/geolis/geolis_link/198801397/ja
36. 藤井直之、**堀井洋一**、宮本正道 (1988), “ユレーライトは直接擬縮物か母天体内沈積物か?--金属粒子の粒形解析から--(演旨)”, 地震学会講演予稿集 1988 巻 2 号 155-155, https://gbank.gsj.jp/geolis/geolis_link/198800574/ja
37. **Youichi Horry** (1990), “What the form tells to us”, Master thesis, Graduate school of Science, Kobe University (1990), 修士論文
38. **Youichi Horii** (1988), “Chondrule formation: Experimental reproduction by Free-fall”, 卒業論文

Patent

1. 堀井洋一, 特願 2015-246901, 事業性分析装置, 2015/12/18
2. 堀井洋一, 特願2015-118276, 水処理システムのシミュレーション方法及びプログラム, 2015/06/11
3. 岸澤夏樹, 堀井洋一, 特願 2014-75215, 逆浸透膜設備設計支援装置及び設備設計支援方法, 2014/04/01
4. 堀井洋一, 都築浩一, 特願 2014-39793, 造水システム, 2014/02/28
5. 堀井洋一, 特願 2013-159503, 事業性評価装置、事業性評価方法及び事業性評価プログラム, 2013/07/31
6. 堀井洋一, 松井恵理, 特願 2013-68753, 逆浸透膜設備性能計算方法、逆浸透膜設備性能計算装置およびプログラム, 2013/03/28
7. 松井恵理, 堀井洋一, 特願 2012-282185, 水質予測装置及び水質予測方法, 2012/12/26
8. 堀井洋一, 手嶋達也, 特願 2014-518084, リスクシナリオ分析システム及びリスクシナリオ分析方法, 2012/05/30
9. 星野剛史, 堀井洋一, 石川忠明, 市毛由希子, 佐藤大樹, 特願 2011-79320, 情報表示端末, 2011/03/31
10. 安藤ハル, 吉村美奈, 堀井洋一, 特願 2011-19635, 組み立て方認識機能を備えた玩具ブロックシステム, 2011/02/01
11. 石川忠明, 堀井洋一, 福間晋一, 特願 2011-1283, アンテナ装置, 2011/01/06
12. 市毛由希子, 堀井洋一, 福間晋一, 特願 2010-76792, 最大酸素摂取量計測装置、最大酸素接種量計測方法及びプログラム, 2010/03/30
13. 荒宏視, 矢野和男, 佐藤信夫, 辻聡美, 脇坂義博, 大久保教夫, 堀井洋一, 特願 2011-507076, コミュニケーション支援装置、コミュニケーション支援システム、及びコミュニケーション支援方法, 2010/03/04
14. 堀井洋一, 福間晋一, 特願2009-282982, 映像表示システム、映像表示方法、およびそれに用いるリモコン装置, 2009/12/14
15. 福間晋一, 堀井洋一, 望月有人, 特願 2009-275828, データ表示装置、データの表示方法及びプログラム, 2009/12/03
16. 石川忠明, 堀井洋一, 石井隆昭, 特願 2009-173276, 車載接続装置、車載接続装置と携帯機器とを接続するコネクタ装置及び車両, 2009/07/24
17. 渡邊純一郎, 堀井洋一, 特願 2009-5925, 検索対象コンテンツの特徴を利用したコンテンツ表示方法, 2009/01/14
18. 大塚理恵子, 堀井洋一, 特願 2008-328005, 複数カメラ制御システム, 2008/12/24
19. 石川忠明, 堀井洋一, 特願 2008-96003, 車載機器の入力装置及び車載機器の入力方法, 2008/04/02
20. 堀井洋一, 特願 2007-335611, コンテンツ配信方法, 2007/12/27

21. 助田浩子, 堀井洋一, 星野剛史, 特願 2007-333441, コンテンツ表示用サーバ, 2007/12/26
22. 助田浩子, 堀井洋一, 特願 2007-301166, コンテンツ格納装置及びコンテンツ格納方法, 2007/11/21
23. 大塚理恵子, 星野剛史, 堀井洋一, 関口茂之, 特願 2007-298295, 表示装置, 2007/11/16
24. 星野剛史, 堀井洋一, 助田浩子, 渡邊純一郎, 特願 2007-247569, リモコンを用いた動画情報処理装置及び動画情報処理システム, 2007/09/25
25. 石川忠明, 大塚理恵子, 堀井洋一, 星野剛史, 特願 2007-129995, 通電クリップ, 2007/05/16
26. 大塚理恵子, 星野剛史, 堀井洋一, 特願 2007-117947, 表示装置, 2007/04/27
27. 堀井洋一, 特願 2007-110766, 操作パネル及び音響発生装置, 2007/04/19
28. 能島雅史, 小村昭義, 板橋武之, 舟越砂穂, 有田裕, 丸山幸伸, 堀井洋一, 石川敬郎, 特願 2007-86333, 燃料環境影響物排出量評価システム, 2007/03/29
29. 廣常朱美, 安齋由美子, 星野剛史, 堀井洋一, 助田浩子, 大塚理恵子, 特願 2007-2540, 情報再生装置、光ディスク及び光ディスクの再生方法, 2007/01/10
30. 助田浩子, 堀井洋一, 特願 2006-282144, 映像記録装置, 2006/10/17
31. 塚田有人, 星野剛史, 松本和己, 堀井洋一, 助田浩子, 特願 2006-239720, コンテンツAND検索方法, 2006/09/05
32. 堀井洋一, 石川忠明, 星野剛史, 中島一州, 丸山幸伸, 特願 2006-239724, 棚状表示機および映像表示方法, 2006/09/05
33. 安齋由美子, 廣常朱美, 向尾将樹, 堀井洋一, 助田浩子, 大塚理恵子, 星野剛史, 特願 2006-160028, 情報記録媒体収納システム及び収納ケース, 2006/06/08
34. 星野剛史, 丸山幸伸, 堀井洋一, 望月有人, 特願 2006-64477, テーブル型情報表示端末, 2006/03/09
35. 大塚理恵子, 堀井洋一, 特願 2005-344875, 自動クリック入力システム, 2005/11/30
36. 渡邊純一郎, 星野剛史, 堀井洋一, セルジオ・パオラントニオ, 特願 2005-331019, リモートコントロール装置, 2005/11/16
37. 石川忠明, 堀井洋一, 星野剛史, 特願 2005-314679, 座標指示入力装置, 2005/10/28
38. 大塚理恵子, 星野剛史, 堀井洋一, 特願 2005-257242, 弾力性素材を用いた入力装置, 2005/09/06
39. 平澤茂樹, 堀井洋一, 助田浩子, 赤司卓也, 星野剛史, 丸山幸伸, 特願 2005-148888, 情報提示装置, 2005/05/23
40. 助田浩子, 水谷世希, 堀井洋一, 星野剛史, 特願 2005-11039, 放送信号の予約録画方法および記録装置, 2005/01/19
41. 尾坂忠史, 一野瀬亮子, 吉田和司, 堀井洋一, 特願 2004-305890, 車両用警告装置, 2004/10/20
42. セルジオ・パオラントニオ, 渡邊純一郎, 星野剛史, 堀井洋一, 特願2004-287246, ダイヤル型操作デバイス, 2004/09/30
43. 大塚理恵子, 星野剛史, 堀井洋一, 長屋茂喜, 特願 2004-267680, 自動追尾機能を

有する撮像装置, 2004/09/15
44. 星野剛史, 大塚理恵子, 堀井洋一, 丸山幸伸, 特願 2005-71604, 立体表示装置, 2004/09/14
45. 星野剛史, 大塚理恵子, 堀井洋一, 丸山幸伸, 特願 2004-267289, 立体表示装置, 2004/09/14
46. 星野剛史, 大塚理恵子, 堀井洋一, 特願2004-266166, 表示装置及びその撮像装置, 2004/09/14
47. 助田浩子, 堀井洋一, 特願 2004-185795, 顧客管理システム, 2004/06/24
48. 井澤真理子, 長屋茂喜, 堀井洋一, 星野剛史, 丸山幸伸, 特願 2004-154075, カーソル制御システム, 2004/05/25
49. 堀井洋一, 木村淳一, 長屋茂喜, 特願 2004-40802, 監視カメラ映像配信システム, 2004/02/18
50. 星野剛史, 堀井洋一, 丸山幸伸, 宮本洋, 加藤真理子, 柳本学, 特願 2007-30693, テーブル型情報端末, 2004/02/13
51. 星野剛史, 堀井洋一, 丸山幸伸, 宮本洋, 加藤真理子, 柳本学, 特願 2007-30696, テーブル型情報端末, 2004/02/13
52. 星野剛史, 堀井洋一, 丸山幸伸, 宮本洋, 加藤真理子, 柳本学, 特願 2004-36745, テーブル型情報端末, 2004/02/13
53. 石川忠明, 星野剛史, 堀井洋一, 特願 2009-63361, 非接触入力インターフェース装置及び情報端末装置, 2003/11/07
54. 石川忠明, 星野剛史, 堀井洋一, 特願2003-378193, 非接触入力インターフェース装置, 2003/11/07
55. 赤城協, 荒井俊史, 堀井洋一, 丸山幸伸, 星野剛史, 特願 2003-322372, 記憶システム, 2003/09/16
56. 助田浩子, 大関一博, 松本千晶, 堀井洋一, 特願 2003-157445, ポイント管理プログラム、該プログラムの提供サーバ及び端末装置, 2003/06/03
57. 星野剛史, 堀井洋一, 特願 2003-76786, 入力表示装置, 2003/03/20
58. 星野剛史, 堀井洋一, 特願 2006-339886, 表示装置, 2003/03/18
59. 星野剛史, 堀井洋一, 特願 2005-68851, 表示装置, 2003/03/18
60. 星野剛史, 堀井洋一, 特願 2003-72917, 商品注文システム, 2003/03/18
61. 星野剛史, 堀井洋一, 特願 2003-73371, 表示装置, 2003/03/18
62. 赤城協, 山口高司, 堀井洋一, 丸山幸伸, 星野剛史, 上船貢記, 荒井俊史, 小田井正樹, 特願 2002-106753, 記憶システムとそれを用いたサービスの販売方法, 2002/04/09
63. 堀井洋一, 丸山幸伸, 中島一州, 星野剛史, 打桐竜巳, 特願 2001-360241, 携帯端末, 2001/11/27
64. 影山昌広, 村上智一, 田邊尚男, 木村淳一, 堀井洋一, 山田佳弘, 柴田明男, 特願 2001-352535, 情報交換方法、情報交換端末装置、情報交換サーバ装置、プログラム, 2001/11/19
65. 鈴木芳典, 堀井洋一, 特願 2001-288008, 動画像符号化・復号化方法, 2001/09/21
66. 星野剛史, 堀井洋一, 丸山幸伸, 特願 2001-252276, 情報表示装置とその操作認識

方法, 2001/08/23
67. 星野剛史, 柴田吉隆, 丸山幸伸, 片山淳詞, 吉丸卓志, <u>堀井洋一</u>, 小嶋聡子, 特願 2001-202376, ネットワーク・システム, 2001/07/03
68. <u>堀井洋一</u>, 特願 2001-174979, 画像作成システム及び画像配信方法, 2001/06/11
69. <u>堀井洋一</u>, 特願 2001-51478, 映像補正方法、映像補正システム、および、映像補正をおこなうためのプログラム記録媒体, 2001/02/27
70. <u>堀井洋一</u>, 丸山幸伸, 星野剛史, 井関大介, 特願 2000-349313, 携帯端末, 2000/11/10
71. <u>堀井洋一</u>, 丸山幸伸, 星野剛史, 中島一州, 特願 2000-349314, 情報検索端末, 2000/11/10
72. <u>堀井洋一</u>, 野村洋之, 特願 2000-17876, 図形作成方法, 2000/01/24
73. <u>堀井洋一</u>, 特願平 10-352259, 3次元輪郭抽出および図形描画方法, 1998/12/11
74. <u>堀井洋一</u>, 特願平 10-352255, 形状分割方法、形状分割装置、形状分割プログラムを記憶した記憶媒体及び形状分割装置作成方法, 1998/12/11
75. <u>堀井洋一</u>, 野村洋之, 特願平 10-352261, 図形作成方法, 1998/12/11
76. <u>堀井洋一</u>, 特願平 10-352258, アニメーション作成方法, 1998/12/11
77. 野村洋之, <u>堀井洋一</u>, 特願平 10-352263, 筆致抽出および図形描画方法, 1998/12/11
78. <u>堀井洋一</u>, 野村洋之, 特願平 10-352262, 図形作成方法, 1998/12/11
79. <u>堀井洋一</u>, 野村洋之, 特願平 10-352260, 図形描画方法, 1998/12/11
80. <u>堀井洋一</u>, 特願平 10-348184, 多層型多角形編集および描画方法, 1998/12/08
81. <u>堀井洋一</u>, 新井清志, 特願平 8-201532, 画像変換方法, 1996/07/31
82. 関本信博, <u>堀井洋一</u>, 特願平 7-315079, 三次元音響提示方法および装置, 1995/12/04
83. <u>堀井洋一</u>, 新井清志, 特願平 7-310319, 透視図作成支援方法, 1995/11/29
84. 新井清志, <u>堀井洋一</u>, 特願平 7-239512, 三次元多関節構造体形状生成方法, 1995/09/19
85. <u>堀井洋一</u>, 藤田実, 大石喜久, 特願平 7-112815, 三次元データ入力システム, 1995/05/11
86. <u>堀井洋一</u>, 特願平 6-221370, 図形描画方法, 1994/09/16
87. <u>堀井洋一</u>, 特願平 6-136845, 音響効果付加方法および装置, 1994/06/20
88. <u>堀井洋一</u>, 特願平 6-135369, 旋律合成方法, 1994/06/17
89. <u>堀井洋一</u>, 安藤ハル, 山口宗明, 特願平 6-59917, 服飾データ管理装置および方法, 1994/03/03
90. 新井清志, <u>堀井洋一</u>, 関本信博, 特願平 6-51340, 透視図作成支援方法および透視図作成支援システム, 1994/02/23
91. <u>堀井洋一</u>, 特願平 5-971, 旋律合成方法, 1993/01/07
92. <u>堀井洋一</u>, 特願平 4-30449, 楽譜入力方法, 1992/02/18
93. <u>堀井洋一</u>, 特願平 4-26507, 音響生成方法, 1992/02/13
94. 菅谷正弘, 伊藤智, 蒲原史朗, 井原茂男, <u>堀井洋一</u>, 特願平 3-281034, 可逆的演算装置, 1991/10/28

95. 伊藤智，蒲原史朗，萱谷正弘，井原茂男，**堀井洋一**，特願平 3-255012，可逆的論理回路および演算装置, 1991/10/02
96. **堀井洋一**，特願平 3-186105，自動伴奏方法, 1991/07/25
97. **堀井洋一**，特願平 3-102476，自動伴奏装置, 1991/05/08
98. ARA KOJI, YANO KAZUO, SATO NOBUO, TSUJI SATOMI, WAKISAKA YOSHIHIRO, OHKUBO NORIO, **HORII YOUICHI**, US13/262673, COMMUNICATION SUPPORT DEVICE, COMMUNICATION SUPPORT SYSTEM, AND COMMUNICATION SUPPORT METHOD, 2010/03/04
99. **HORII YOUICHI**, WO2016JP081354W, FEASIBILITY ANALYSIS DEVICE, 2016/10/21
100. **HORII YOUICHI**, TSUZUKI KOICHI, CN201510089067A, FRESH WATER GENERATION SYSTEM, 2015/02/27
101. KISHIZAWA NATSUKI, **HORII YOUICHI**, WO2015JP051571W, DEVICE FOR SUPPORTING DESIGN OF REVERSE OSMOSIS MEMBRANE EQUIPMENT, AND METHOD OF SUPPORTING DESIGN OF EQUIPMENT, 2015/01/21
102. **HORII YOUICHI**, US14/445092, BUSINESS VIABILITY EVALUATION APPARATUS, BUSINESS VIABILITY EVALUATION METHOD, AND BUSINESS VIABILITY EVALUATION PROGRAM, 2014/07/29
103. **HORII YOUICHI**, MATSUI ERI, US14/655001, REVERSE OSMOSIS EQUIPMENT PERFORMANCE CALCULATION METHOD, REVERSE OSMOSIS EQUIPMENT PERFORMANCE CALCULATION DEVICE AND PROGRAM, 2013/12/27
104. **HORII, YOUICHI**, TESHIMA, TATSUYA, WO2012JP003519W, RISK-SCENARIO ANALYSIS SYSTEM AND RISK-SCENARIO ANALYSIS METHOD, 2012/05/30
105. HOSHINO TAKESHI, **HORII YOUICHI**, ISHIKAWA TADAAKI, ICHIGE YUKIKO, SATO HIROKI, US13/397065, INFORMATION DISPLAY TERMINAL, 2012/02/15
106. ISHIKAWA, TADAAKI, **HORII, YOUICHI**, MARUYAMA, YUKINOBU, SHIGENARI, YUKIO, WO2011JP003694W, APPARATUS AND SYSTEM FOR SUPPLYING OPTICAL ENERGY, 2011/06/29
107. SUKEDA HIROKO, **HORII YOUICHI**, HOSHINO TAKESHI, US12/314887, SERVER FOR DISPLAYING CONTENTS, 2008/12/18
108. OTSUKA, RIEKO, HOSHINO, TAKESHI, **HORII, YOUICHI**, EP08007542A, STEREOSCOPIC PROJECTION DISPLAY APPARATUS, 2008/04/17
109. NOJIMA MASAFUMI, KOMURA AKIYOSHI, ITABASHI TAKEYUKI, FUNAKOSHI SUNAO, ARITA HIROSHI, MARUYAMA YUKINOBU, **HORII YOUICHI**, ISHIKAWA TAKAO, US12/034180, EVALUATION SYSTEM FOR AMOUNT OF EMISSION GASES THROUGH FUEL SUPPLY CHAIN, 2008/02/20
110. **HORII YOUICHI**, ISHIKAWA TADAAKI, HOSHINO TAKESHI, NAKAJIMA ISSHU, MARUYAMA YUKINOBU, US11/822957, SHELF-LIKE

DISPLAY MACHINE AND IMAGE DISPLAY METHOD, 2007/07/11

111. YOSHINAGA TOMOAKI, NAGAYA SHIGEKI, **HORII YOUICHI**, US11/737159, ATTENTION LEVEL MEASURING APPARATUS AND AN ATTENTION LEVEL MEASURING SYSTEM, 2007/04/19
112. ANZAI YUMIKO, HIROTSUNE AKEMI, MUKOH MASAKI, **HORII YOUICHI**, SUKEDA HIROKO, OTSUKA RIEKO, HOSHINO TAKESHI, US11/696284, INFORMATION RECORDING MEDIUM STORING SYSTEM AND STORING CASE, 2007/04/04
113. OTSUKA RIEKO, HOSHINO TAKESHI, **HORII YOUICHI**, US11/483159, INPUT DEVICE USING ELASTIC MATERIAL, 2006/07/10
114. OSAKA TADASHI, ICHINOSE RYOKO, YOSHIDA KAZUSHI, **HORII YOUICHI**, US11/252582, WARNING DEVICE FOR VEHICLES, 2005/10/19
115. SUKEDA, HIROKO, **HORII, YOUICHI**, EP05006411A, CUSTOMER INFORMATION MANAGEMENT SYSTEM, 2005/03/23
116. HOSHINO TAKESHI, OTSUKA RIEKO, **HORII YOUICHI**, MARUYAMA YUKINOBU, US11/072495, STEREOSCOPIC DISPLAY APPARATUS, 2005/03/07
117. HOSHINO TAKESHI, **HORII YOUICHI**, MARUYAMA YUKINOBU, MIYAMOTO YOH, KATO MARIKO, YANAGIMOTO MANABU, CN200510007575A, TABLE TYPE INFORMATION TERMINAL, 2005/02/05
118. **HORII YOUICHI**, KIMURA JUNICHI, NAGAYA SHIGEKI, US11/037337, IMAGE DISTRIBUTION SYSTEM OF SURVEILLANCE CAMERA WITH PRIVACY PROTECTION, 2005/01/19
119. OTSUKA RIEKO, HOSHINO TAKESHI, **HORII YOUICHI**, US11/630152, DISPLAY SYSTEM AND CAMERA SYSTEM, 2004/09/10
120. HOSHINO TAKESHI, OTSUKA RIEKO, ISHIBASHI ATSUSHI, **HORII YOUICHI**, YANAGIMOTO MANABU, MARUYAMA YUKINOBU, CN200410074819A, DISPLAY APPARATUS AND IMAGE PICKUP APPARATUS, 2004/08/30
121. HOSHINO TAKESHI, **HORII YOUICHI**, OTSUKA RIEKO, US10/785051, DISPLAY APPARATUS, 2004/02/25
122. **HORII YOUICHI**, MATSUMOTO, CHIAKI, OHZEKI, KAZUHIRO, SUKEDA, HIROKO, KR20030059857A, POINT MANAGING PROGRAM USING MOBILE PHONE, SERVER FOR OFFERING THE SAME, AND TERMINAL DEVICE, 2003/08/28
123. HOSHINO TAKESHI, SHIBATA YOSHITAKA, MARUYAMA YUKINOBU, KATAYAMA ATSUSHI, YOSHIMARU TAKASHI, **HORII YOUICHI**, KOJIMA SATOKO, US10/082107, REMOTE CONTROL NETWORK SYSTEM FOR REMOTELY CONTROLLING A PLURALITY OF ELECTRIC APPARATUS SUCH AS HOME APPLIANCES, 2002/02/26
124. **HORII YOUICHI**, HOSHINO TAKESHI, ISEKI DAISUKE, MARUYAMA YUKINOBU, KR20010046728A, PORTABLE TERMINAL, 2001/08/02

125. **<u>HORII</u> <u>YOUICHI</u>**,　NOMURA HIROYUKI, US09/514283, METHOD OF DRAWING SHAPES, 2000/02/28
126. **<u>HORII</u> <u>YOUICHI</u>**, US09/455405, METHOD OF CREATING TWO DIMENSIONAL SHAPES DRAWING SYSTEM, 1999/12/06
127. **<u>HORII YOUICHI</u>**,　ARAI,　KIYOSHI, US08/904449, PICTURE CONVERSION METHOD AND MEDIUM USED THEREFOR, 1997/07/31
128. **<u>HORII YOUICHI</u>**, US08/306153, METHOD FOR THREE-DIMENSIONALLY DRAWING FIGURE ON DISPLAY PLANE, 1994/09/14
129. **<u>HORII YOUICHI</u>**, US08/302441, METHOD OF COMPUTER MELODY SYNTHESIS RESPONSIVE TO MOTION OF DISPLAYED FIGURES, 1994/09/09

Educational activities

1. 女子美術大学　メディア文化論特講「企業におけるグラフィクスとインタラクションの役割とワークスタイル」, 2014/6/26
2. 立命館大　キャリア形成科目 「技術者のキャリア」リレー講座「日立ヒューマンインタラクションラボ研究の取り組み」, 2004/12/7
3. 新潟大学　集中講義「情報メディア特殊研究Ｉ」, 2000/9/11 – 9/14

Awards

1. 計測自動制御学会　優秀l講演賞, “逆ベイズ推定を用いた意思決定のモデル化”, 2016/12/15
2. IEEE Consumer Electronics Society, Chester W Sall Memorial Awards, ”Information-Accessing Furniture to Make Our Everyday Lives More Comfortable”, 2008/1/12