IIAI Open Conference Publication Series
IIAI Letters on Business and Decision Science
Vol. 002, LBDS061
DOI: https://doi.org/10.52731/lbds.002.061

A Study on Artificial Intelligence and Human Decision-Making

Takaaki Hosoda *

Abstract

This paper focuses on Explainable AI, which has become a hot topic in recent years, and discusses the possibility of a decision-making scheme in which artificial intelligence decision making and human decision making are linked, focusing on the human condition. In examining the relationship between AI and decision makers, the paper argues that the decision maker's own sense of responsibility is important for the decision maker to accept the AI's reasoning results.

Keywords: Acceptability, Explainability, Decision making process, AI, Responsibility

1 Introduction

In recent years, there have been remarkable technological advances related to artificial intelligence (Hereafter, AI). The introduction of technologies related to AI is progressing rapidly in a variety of fields. For example, AI is now taking the place of human intuition and experience in such areas as automated driving, automatic diagnosis in medicine, and AI investment. Such developments have led to the realization of various services that utilize AI, and we make decisions on a daily basis while making use of these services.

On the other hand, it is not necessarily true that we can rely on the decisions made by AI. For example, even if a person in charge of practical business decisions in a company explains that he or she will choose an alternative because it is a choice made by AI, his or her superiors, managers and executives, are unlikely to accept the choice. In the case of life-threatening medical diagnoses, if there is no basis for the diagnosis, the patient will feel anxious, and it is imaginable that the patient will not be able to make a decision based on the results alone.

Thus, while the development of AI can be very useful for human decision making, there are also aspects of decision making in which we cannot accept AI, which determines solutions in a black box manner. In response to this problem, Explainable AI (Hereafter, XAI) has emerged as a technology that outputs additional information (interpretation of the model, explanation of the basis for the decision, etc.) in addition to the results output by the learning model that constitutes the AI to assist its output. XAI has gained significant attention since around 2018, when technologies related to AI were increasingly being used in corporate activities in Japan. Along with the growing use of AI technologies, the reason for the increased attention was the strong interest and concern about the social issues that would be brought about in the future. In response to this

* Advanced Institute of Industrial Technology, Tokyo, Japan

growing social interest, various organizations, especially government agencies, have taken up XAI as one of their important research areas.

This is also evident in surveys. According to a survey by the Information-technology Promotion Agency, Japan (IPA), based on a questionnaire to companies, the challenges of social implementation of AI are listed as "development challenges such as engineer shortage," "challenges of characteristics of AI such as explainability and safety," "challenges of international competitiveness compared to the United States and China," "challenges related to legal systems," "challenges related to AI It is evident that entrepreneurs also understand the importance of XAI.[1] [2] As can be seen from this survey, although AI technology is attracting attention, in reality, there are social or technological barriers to the social implementation of AI. This paper focuses on Explainable AI, which has been a hot topic in recent years, and discusses the possibility of a decision-making scheme in which decision-making by AI and human decision-making work together, focusing on the state of the human decision-making process that accepts the decision-making solution derived by AI.

2 Explainable AI

Explainability refers to the high ability of an AI to explain "why it gave that answer" for the answers it derives and refers to AI with a high level of this ability. According to B. Arrieta et al [3], "In order to avoid limiting the effectiveness of the current generation of AI systems, Explainable AI (XAI) proposes creating a suite of ML techniques that 1) produce more explainable models while maintaining a high level of learning performance (e.g., prediction accuracy), and 2) enable humans to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners. XAI draws insights as well from the Social Sciences and considers the psychology of explanation."

For example, by having the AI learn by reading many normal and abnormal images in CT images of cancer patients, the AI can determine the presence or absence of cancer by analyzing the CT images. In doing so, the XAI should be able to explain what features in the image it focused on to reach that conclusion. In this way, the ability to explain the basis for the AI's inference results, and for that explanation to be close to the criteria for human judgment, is a necessary requirement for the AI to be able to replace human work. The company has developed a system that can do this with a high degree of accuracy [4]. This feature explains the presence of abnormalities by indicating that the red-colored areas in the CT image of the eye (left in the figure) and the AI's description (middle and right in the figure) are the diseased areas.

Thus, the attention to Explainable AI has increased rapidly in the past few years because it provides a rationale for the choices that the AI derives [3].

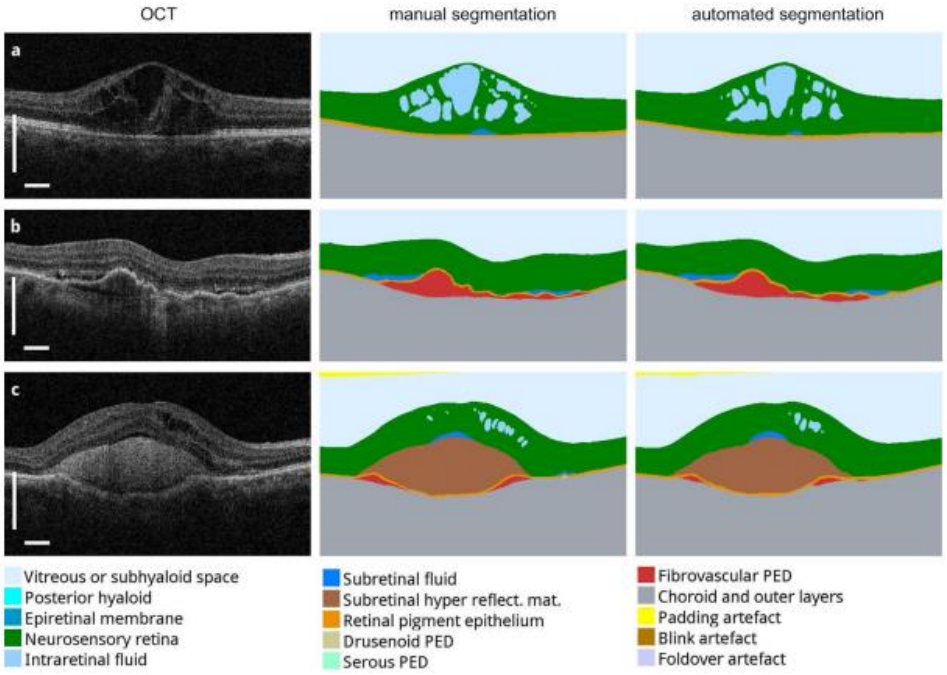

Figure 1: Results of the segmentation network [3]

In the background, due to the characteristics of the models composed by complex neural networks, the internal logic is a black box compared to conventional machine learning methods, making it difficult to explain the inference process. This has led to the recent active research of XAI methods that can be applied to machine learning methods, particularly deep learning. This is reflected in the number of publications related to Explainable AI, and the trend is also on the rise [3].

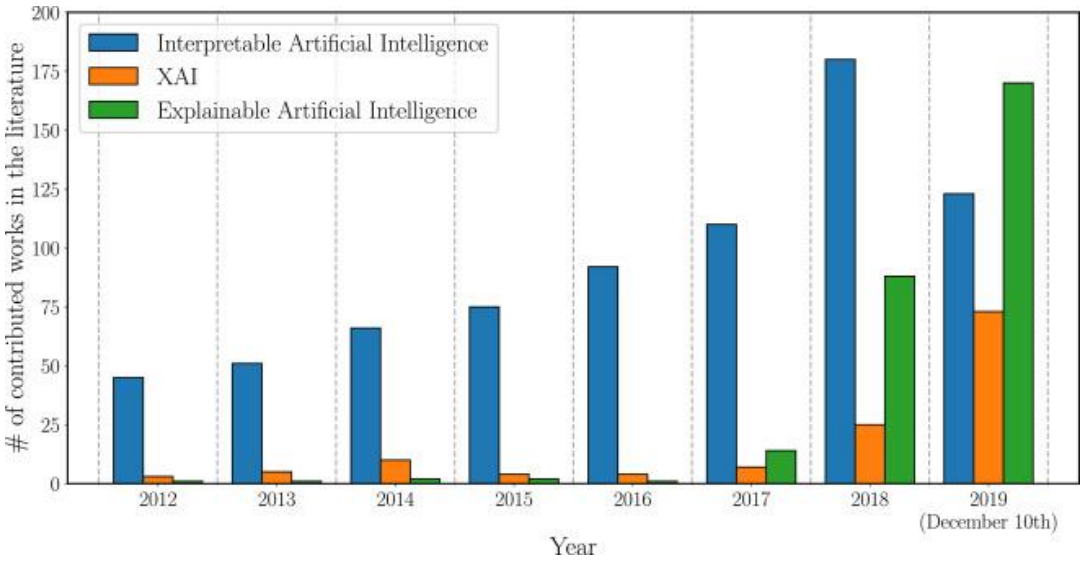

Figure 2: Evolution of the number of total publications whose title, abstract and/or keywords refer to the field of XAI.

3 What it Takes to Accept

The previous section has shown that Explainable AI requires decision makers to accept the AI's reasoning by showing them the basis for the AI's reasoning results, which until now could not be shown. The decision maker is required to accept the AI's inference results.

So, what kind of decision-making process do humans who accept AI's inference results go through in making their decisions? What do they value in their decision-making process? In this section, some studies on the human decision-making process will be reviewed.

3.1 The Decision-Making Process in Simon's Satisficing Principle

Simon argued that humans under Limited Rationality make decisions according to the Satisficing Principle, which states that humans set a certain goal level and choose an alternative when they discover an alternative that can achieve that goal level [5]. Simon further analyzed the decision-making process under limited rationality by dividing the decision-making process into four activities: intelligence activity, design activity, selection activity, and reconsideration activity. Figure 3 is a schematic representation of these activities [6].

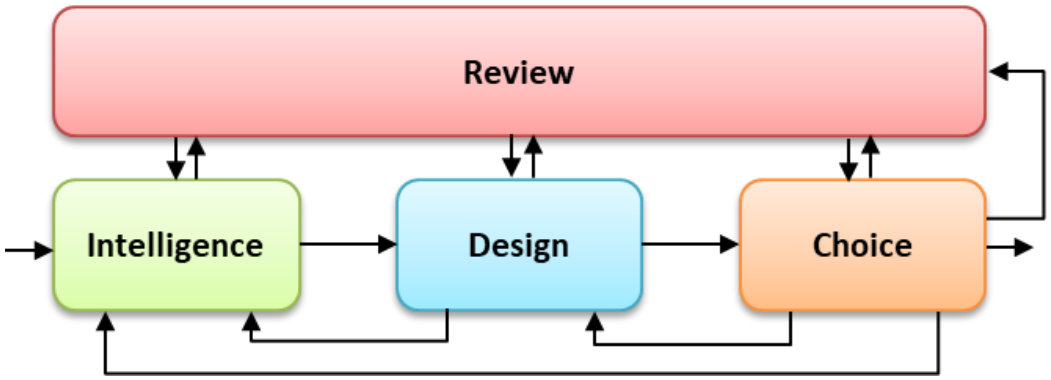

Figure 3: Flow of decision-making process in Simon (prepared by the author) [6]

Simon further divides the decision-making process into several activity processes and analyzes in detail the procedures by which human beings make decisions under limited rationality.
In The new science of management decision published in 1960, he developed the idea that the decision-making process involves three activities: intelligence activity, design activity, and selection activity.

Information activity is the activity of clarifying the decision-making problem by collecting information for decision-making and clarifying the problem. Information activities are the beginning of the decision-making process, and the fact that the decision-maker is able to recognize a problem indicates that there is some gap between the reality and the ideal.

Design activities are activities to discover, develop, and explore possible alternatives to the problems identified through information activities. Discovery and development can be considered to include the consideration of alternatives based on past experiences and actions.

Selection activities are activities to select alternatives derived in the design activities for the decision-making problem set in the information activities. As the word "selecting" is used in the original text, it is not only a selection activity but also an activity that includes evaluation and examination of alternatives.

Information, design, and selection activities occur in this order. Simon, by the way, added a fourth activity, reconsideration activity, to the 1977 revision of the previous document. Reexamination activity is an activity to review and evaluate previous decisions. This activity generates feedback among the information, design, and selection activities. In other words, Simon argues that the decision-making process generally proceeds in a bidirectional and interactive manner, rather than in a unidirectional progression of each activity.

Based on Simon's assertion above, the author believes that Simon's three-phase part is a flow diagram as shown in Fig3.

3.2 Montgomery's Dominance Search Model of the Decision-Making Process

Based on his observations of human decision-making behavior, Montgomery found that humans generally make decisions based on heuristics rather than the expected utility maximization principle [7]. Montgomery asserted that there are three criteria for evaluating heuristic-based decision-making rules: complete applicability (one and only one alternative can be chosen), ease of justification (the reasons for the choice can be explained and defended to oneself and others), and ease of application (ease of application to actual decision-making problems). If the only solution cannot be found, another rule that replaces the superiority rule is successively applied to transform the decision-making problem into a structure in which the superiority rule can be applied. Montgomery named this type of decision-making process the dominant search model. The Dominance Search Model is a flowchart like the one shown in Fig4, and it goes through many feedback loops before the process is completed. Through these feedback loops, people reflect on their own decision-making process. By doing so, humans discover a problem structure that allows them to justify their decisions, and thus they can clearly assert the reasoning that justifies the only solution that can be derived from it. Montgomery argues that humans make decisions through such a decision-making process.

3.3 The Decision-Making Process in the Seeking Model for Acceptable Decision-Making Process

The author's survey of Simon and Montgomery's previous research and the author's own practical experience suggest that humans often make decisions by reflecting on their own decision-making process and choosing alternatives obtained when they are satisfied with the process. Furthermore, the author found that when people are convinced by the decision-making process, they gain confidence and rationale to justify their own decisions. Based on these suggestions, the author came up with the concept of "acceptable decision-making". The original meaning of the word "acceptability" is "to understand and accept the ideas and actions of others as valid. In other words, the concept of conviction is essentially a concept that indicates a judgment about another person (object). For this reason, the author would like to define "satisfaction with the decision-making process" as "the objectification and evaluation of one's own decision-making pro-

cess, and acceptance of it in light of one's own sense of values. Therefore, an acceptable decision is a decision to choose an alternative obtained at the point of being satisfied with one's decision-making process in such a sense.

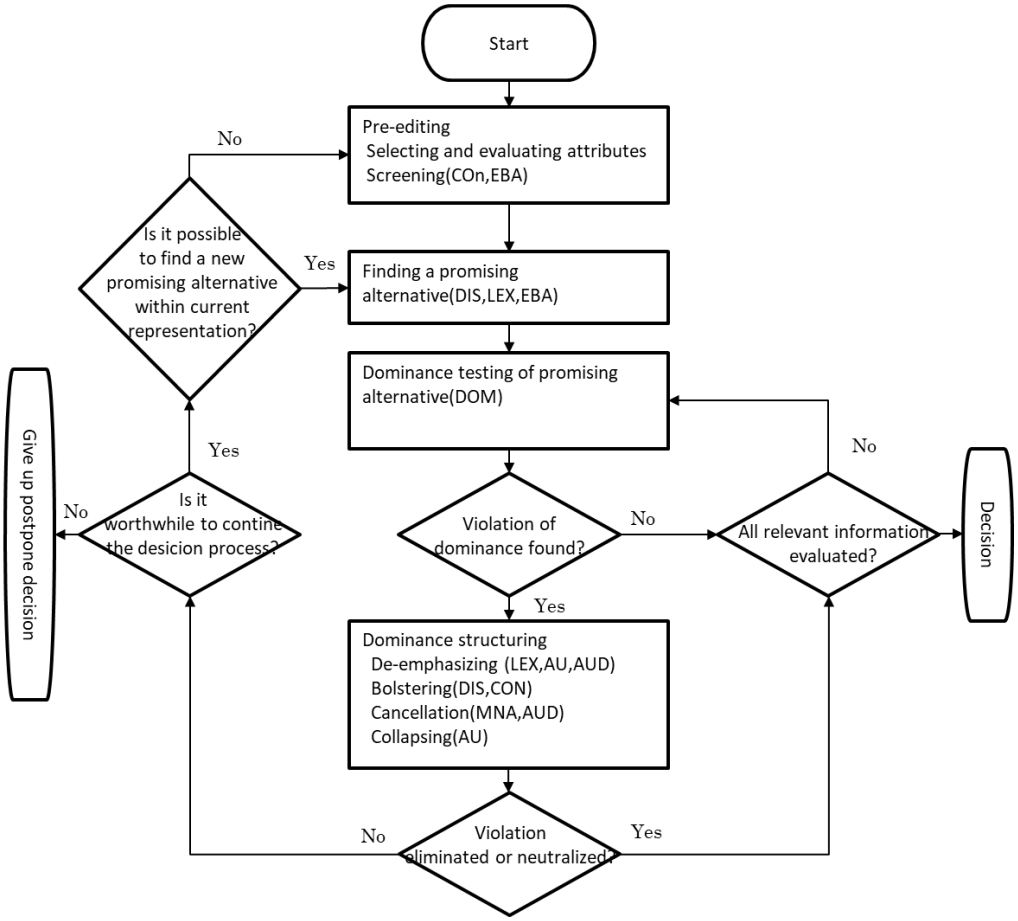

Figure 4: Montgomery's Dominance Search Model [7]

The conceptual model of acceptable decision-making introduced by the author is shown in Fig 5. Based on this conceptual model, I would like to propose a practical method that can realize acceptable decision-making as the "acceptable process seeking model. The acceptable process seeking model consists of the flowchart shown in Figure 6.

The process flow shown in Figure xx illustrates the conceptual model of Acceptable Decision Making introduced by Hosoda [8][9]. First, "candidate consideration and derivation" is the process of understanding the relationship between attributes and alternatives. This understanding provides an objective understanding of the trade-off relationship between alternatives and their attributes in a decision-making problem. Next, "Review of Problem Structure" is the process of reviewing the problem structure identified in the "Consideration and Derivation of Alternatives" process. Through this review process, the understanding of the problem structure can be more acceptable when narrowing down the options to be selected. Finally, the "review for acceptability structure" is the process of objectifying the decision-making process followed by the decision-maker and reviewing whether the process is acceptable in light of one's values, outlook on

life, and worldview. This review for introspection can further improve the sense of acceptability The process of the acceptable process seeking model starts with selecting attributes and applying the superiority rule on the attributes using the concept of the dictionary compilation rule, which derives alternatives by considering the most important attributes first, and then applying the superiority rule to derive alternatives as candidates while trying to increase or decrease the number of attributes to be considered, if necessary. Through this decision-making process, the structure of the decision problem is grasped, and an alternative is selected after finding a acceptable reason for selecting that alternative. The Acceptable Process Seeking Model, so to speak, is an approach to decision making in which the decision is not based on the solution, but on the process of selection leading to the solution.

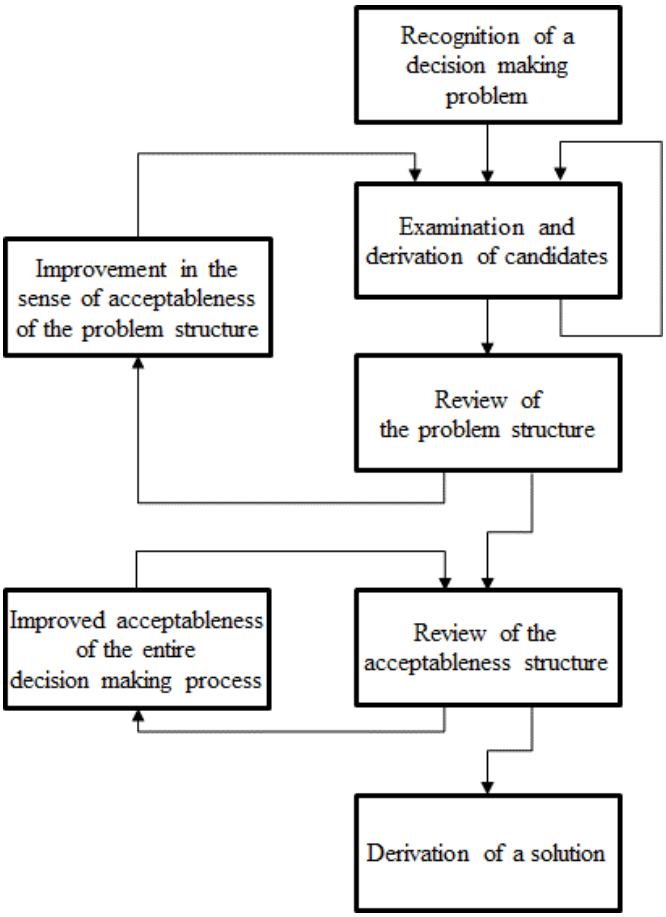

Figure 5: Conceptual Model of Acceptable Decision Making [8][9]

4 Discussion on the Relationship between AI and Human Decision Making

XAI attempts to provide a basis for solution derivation by showing the decision maker the process of deriving inference results. However, as previous studies have shown, the decision-making process plays an important role in human decision making. In the decision-making

models of Simon, Montgomery, and the author, there is a feedback loop, and the structure requires deliberation in the process of going through the feedback loop.

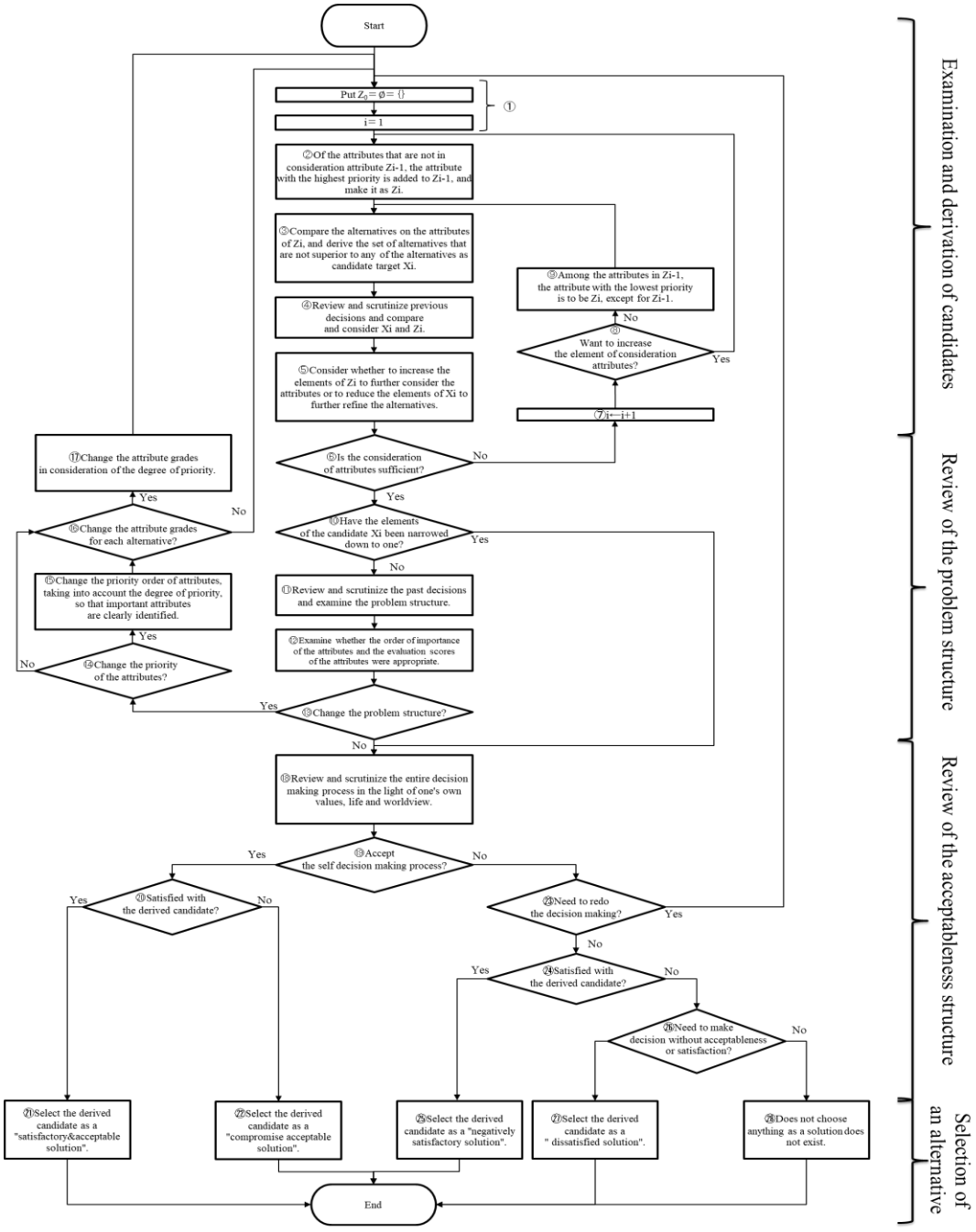

Figure 6: The Acceptable Process Seeking Model [8][9]

Among them, Montgomery's Dominance Search Model requires a justification argument in the decision-making process, while the author's Acceptable Decision-Making model requires the decision-maker to be convinced of the choice process leading to a solution. Thus, observing the state of the human decision-making process suggests that simply explaining the results of reasoning may not be sufficient.

Simon states that human rationality is limited, but when the process is considered, it becomes procedurally rational. There needs to be a process leading to the rationale of what reasoning process led to the reasoning result. The same is equally true for Montgomery's argument for justified decision-making. In order to claim legitimacy, it is necessary to persuade other companies of the rationale for choosing that option based on the strength of the rationale and the order of precedence. This will be difficult to do with only a candidate choice and the rationale for the choice.

Finally, it would be even more difficult in the case of the Acceptable Decision-Making model, which the author argues the idea is that humans become more receptive through their understanding of the decision problem and their understanding of the decision-making process. This is fostered as a result of proactively working on the decision-making problem. Therefore, it is difficult for humans to proactively accept decision-making by AI.

How, then, can humans utilize AI with an emphasis on the decision-making process? One tentative approach is to appeal to the decision-maker's sense of responsibility. No matter how accurate the reasoning results and rationale are, if the accepting decision maker uses the solution without a sufficient sense of responsibility, the AI will not be fully satisfactory after implementing the reasoning result and the decision maker will not be fully satisfied. AI will not be fully satisfactory after the inference results are implemented. Therefore, the author believes that it is important to cultivate a sense of responsibility among decision makers as part of decision support.

AI does not solve all problems by itself. In many cases, AI can replace human tasks with greater precision. However, as far as atypical decisions are concerned, the decision makers are the ones who make the decisions using AI; it is the decision makers who ultimately determine whether or not they have confidence in the AI's conclusions, and it is they who are responsible for those conclusions that establish the relationship between AI and the decision makers. In the future, it will be commonplace for AI to be able to make highly accurate inferences and to show acceptable results, but the author believes that the probability of a decision-making scheme in which AI and human decision makers work together is important. The author believes that the probability of a decision-making scheme in which AI and human decision-making are linked is important.

5 Conclusion and Future Work

In order to discuss the relationship between AI and decision makers, this paper focuses on Explainable AI and discusses the possibility of a decision-making scheme in which decision-making by AI and human decision-making are linked, focusing on the state of human decision-making processes that accept the decision-making solutions derived by AI.

While the development of AI is expected to continue in the future, further research on human decision-making using AI will be necessary. The author would like to continue to work on the state of decision makers using AI as a research topic in the future.

References

[1] "AI White Paper 2019", Information-technology Promotion Agency, Japan, 2019; https://www.ipa.go.jp/ikc/info/20181030.html

[2] "Survey Report on the Promotion of AI Social Implementation", Information-technology Promotion Agency, Japan, 2019; https://www.ipa.go.jp/files/000067229.pdf

[3] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, *58*, 82-115.

[4] De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., ... & Ronneberger, O. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine*, *24*(9), 1342-1350.

[5] Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). MIT press.

[6] Simon, H. A. (1960). The new science of management decision.

[7] Montgomery, H. (1983). Decision rules and the search for a dominance structure: Towards a process model of decision making. In Advances in psychology (Vol. 14, pp. 343-369). North-Holland

[8] HOSODA, T. (2017). A Study on Methods of Practical Decision MakingConstruction of a Conceptual Model for Acceptable Decision Making-, Tokyo, Waseda University, Ph. D. thesis

[9] Hosoda, T., & Maruyama, H. (2022). Differences between the Purposes of the Dominance Search Model and the Acceptable Decision-Making Search Model. *International Journal of Service and Knowledge Management*, *6*(2).