Comparison of Condition Parts of Decision Rules and Optimum Solutions by Using Identity Mapping Model in Consideration of Criteria of Preference

Yusuke Enomoto
Graduated School of Wakayama University
usuke@super-r.net

Toshinobu Harada
Wakayama University
harada@sys.wakayama-u.ac.jp

930 Sakaedani, Wakayama-city Wakayama 640-8450 Japan

Abstract—**The rough sets approach has attracted attention as a method of knowledge acquisition. However, few studies verify availability by comparing rough sets with methods other than C4.5. Therefore, in this study, we applied rough sets and an inverse reasoning system with a five layer neural network as methods for solving inverse problems. The aim of this study is to clarify the features of the condition parts of decision rules and optimum solutions. Concretely, we compared condition parts of decision rules with a high covering index value as obtained with optimum solutions that were found by exploring the entire range of solutions space. In addition we considered the feature of both methods, and were able to clarify the features of each.**

1. INTRODUCTION

When a new product is planned, the relationship between the attributes (cause) of the planned product and a specific evaluation (result) is obtained by using data taken from a survey. This is a problem that explores a cause from a result, and that is defined as an inverse problem. To solve inverse problems is generally difficult. When an inverse problem is to be solved with methods such as neural networks and genetic algorithms, more than one optimum solution can be obtained.

However, even after one or more optimum solutions are obtained, researchers may wonder if those solutions are the only possible ones.

Generally, in inverse problems, more than one combination of attributes can satisfy a particular evaluation, and many combinations exist. That is to say, researchers can only solve an inverse problem completely after the entire range of the possible solutions space has been grasped, and they have obtained all the possible solutions that can satisfy a particular evaluation. Only then, can one be sure one has solved an inverse problem completely.

Therefore, in our study, we applied rough set theory and an inverse reasoning system with a five layer neural network as methods for solving an inverse problem. The neural network is based on an identity mapping model (hereafter, referred to as the identity mapping model). The rough sets approach has attracted attention as a method of knowledge acquisition [1]. However, few studies verify the availability by comparing rough sets with methods other than C4.5 [2, 3]. Therefore, the aim of this study is to clarify the features of the condition parts of decision rules and optimum solutions.

Table. 1 Item and category

Item	Category	Rough set theory	Identity mapping model
Price	0~30,000 yen	a1	0~0.012
	30,001~40,00 yen	a2	0.012~0.115
	40,001~50,000 yen	a3	0.115~0.218
	50,001~100,000 yen	a4	0.218~0.732
	100,001 yen ~	a5	0.732~1
Maximum output	0~30W	b1	0~0.174
	31~50W	b2	0.183~0.357
	51W~	b3	0.366~1
Sound functions	1	c1	0
	2	c2	0.33
	3	c3	0.67
	4	c4	1
MDLP	not equipped	d1	0
	equipped	d2	1
Recording speed of MD	not equipped	e1	0
	double	e2	0.33
	quadruple	e3	1
MD changer	not equipped	f1	0
	equipped	f2	1
CD changer	not equipped	g1	0
	3 CDs	g2	0.5
	5 or more CDs	g3	1
Cassette deck	not equipped	h1	0
	equipped	h2	1
Audio devices	1-2	i1	0~0.2
	3	i2	0.4
	4	i3	0.6
	5	i4	0.8
	6 or more	i5	1

First, we adopted consumer preferences regarding audio products as our case study and conducted a preference survey of 40 subjects. Second, we calculated the condition parts of decision rules (hereafter, DRs), and optimum solutions that satisfy each subject's preferences by applying rough set theory and an identity mapping model to the results of the survey. Third, we compared the optimum solutions obtained by exploring the entire range of solutions space with DRs that were obtained by using rough set theory. We examined the range of optimum solutions by using the identity mapping model and those DRs with a high covering index (hereafter, C.I.) value. Moreover, we searched for the categories to which survey subjects attached importance when they purchased audio products. Then, we compared these

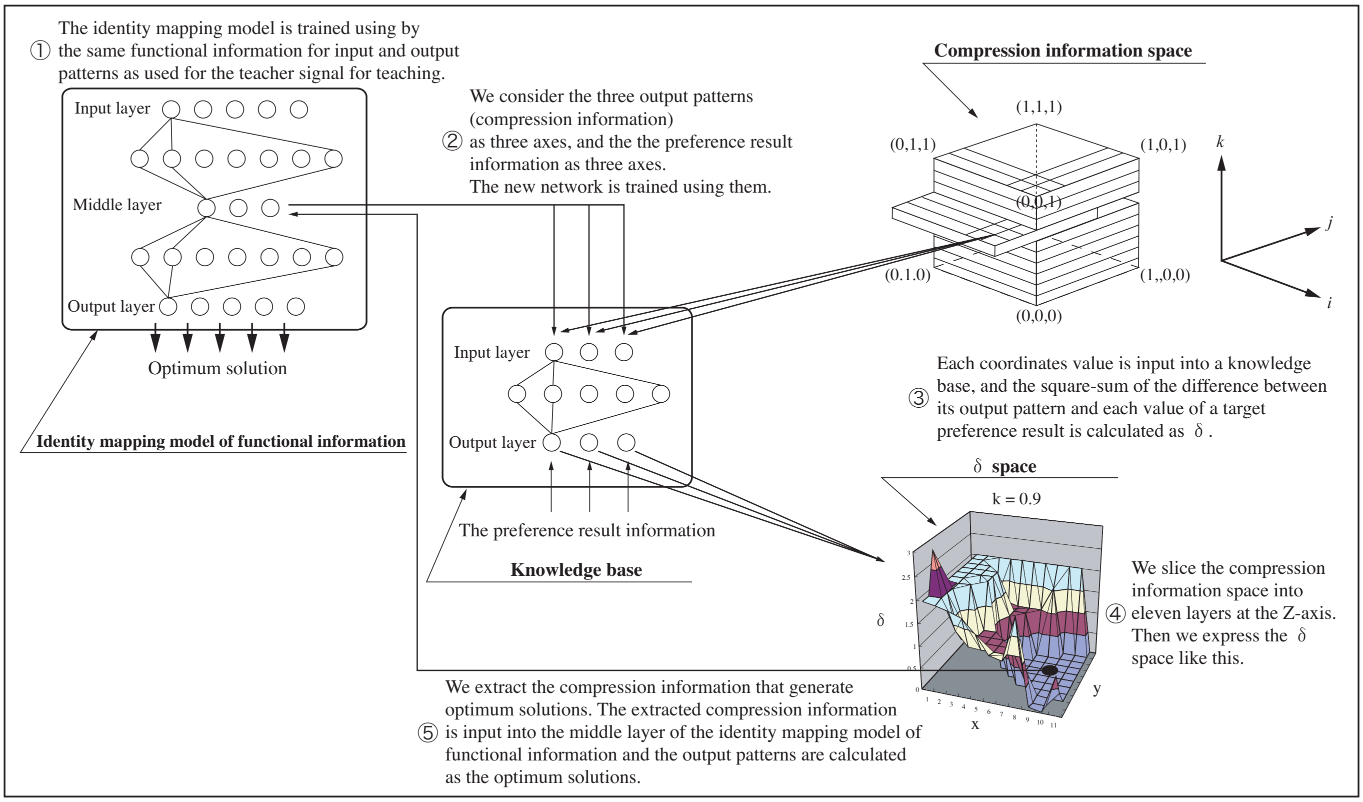


Figure. 1 Search of various solutions using inverse reasoning system

categories with the obtained DRs and optimum solutions.

2. CATEGORY CLASSIFICATION BY PREFERENCE SURVEY OF AUDIO PRODUCTS

We surveyed 40 subjects and then classified the price, the functions, and the number of audio devices that they owned into nine items and 29 categories [4]. Table 1 shows these items and categories. Then, in a preference survey, we showed the 40 subjects the shape and functions of 47 marketed products of nine audio makers, and had these subjects give their preference for each audio product at three levels; "I want to purchase", "I can not decide whether I would purchase", and "I do not wish to purchase". In addition, we had 18 subjects suppose that they purchased audio products. And we had them choose categories of five higher ranks that they considered important from Table 1. Hereafter, we call these factors criteria of preference.

3. INVERSE REASONING SYSTEM WITH IDENTITY MAPPING MODEL

3.1 Identity Mapping Model

A Neural Network model is classified into type according to the training pattern and form of the network, e.g., perceptron type or Hopfield. The features of an identity mapping model are a symmetrical network form in the middle layer, and the use of the same signal for input and output patterns as is used for the teacher signal for training the model [5].

The identity mapping model is used as follows. First, the network is separated into two networks with a hidden layer located in the middle (hereafter, middle layer) after training the model. Next, a specific pattern is input and the output pattern of the middle layer is used as a compression form of the input pattern. Then, the compression information is arbitrarily extracted and used to generate an optimum solution. In the final step, the compression information that generates the optimum solutions is used as the pattern input into the middle layer (input layer after separation), and the resulting output pattern is calculated as optimum solutions.

3.2 Extraction of Optimum Solutions by Using Identity Mapping Model

Figure 1 shows an explanatory diagram of the inverse reasoning system with the identity mapping model used in this study. Details of its construction follow.

3.2.1 Construction of identity mapping model

Using the functions of the audio products (hereafter, functional information) the identity mapping model is trained. In this study, we set the numbers of units for both an input layer and an output layer to nine, because the functional information consisted of nine items. Then we selected the numbers of units for the second, middle, and fourth layers to (15, 3, 15) as the set having fewest errors after trying several combinations of numbers of units. In addition, we set the number for the middle layer to three units because we wished to consider three values of output patterns of that layer as the values of the X, Y, and Z axes of a three-dimensional space (hereafter, 3D space). Here, we normalized raw data so that the maximum value of the function of each item became 1 and the minimum value became 0, and we adopted these data

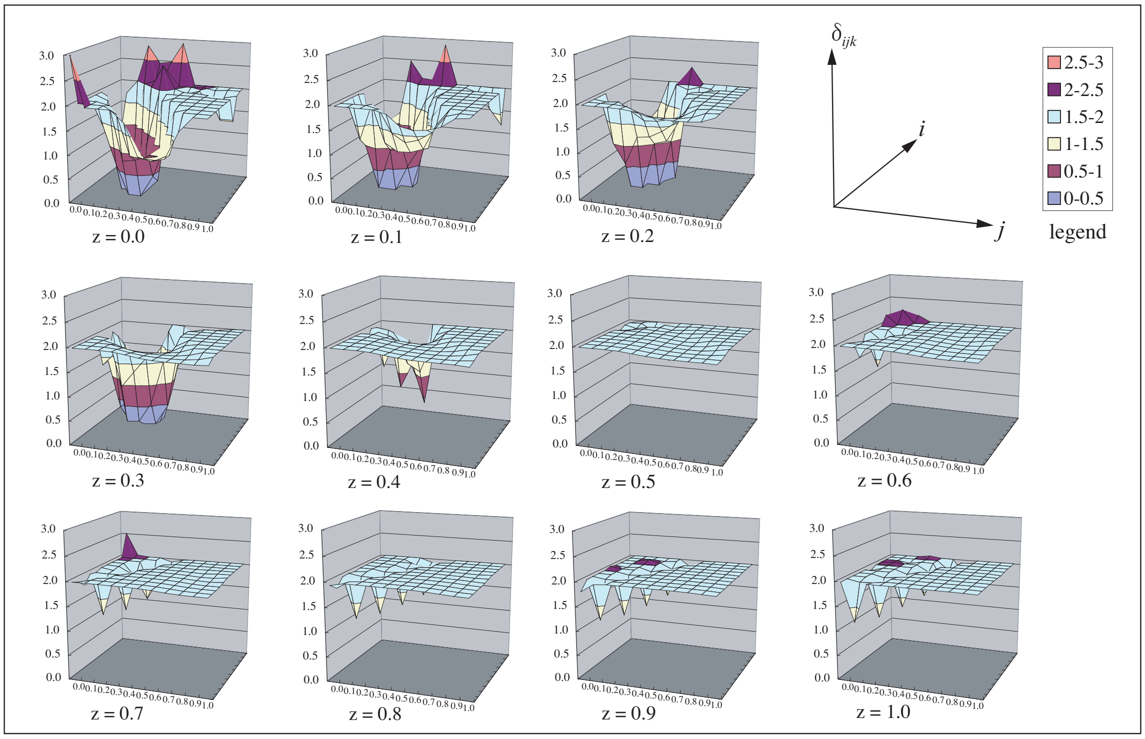


Figure. 2 δ space of a certain subject, A

Table. 2 Correlation coefficient

Optimum solutions - CPs	CPs - Preference samples	Optimum solutions - Preference samples
0.709	0.773	0.859

as the teacher signals for training the model (see Table 1).

3.2.2 Neural network training of functional information and preference results

After the training process, the identity mapping model of functional information is separated at the middle layer. Next, the input pattern of each sample is input into this network, and the output pattern in the middle layer corresponding to each input pattern can be calculated. We call this output pattern compression information. Furthermore, a new network is defined, and the network is trained by the teacher signal that uses the compression information from each sample as input patterns, and uses a preference result [wanting to purchase →(0,0,1), can not decide whether→(0,1,0), not wishing to purchase→(1,0,0)] as output patterns. We call this network the knowledge base. We created this knowledge base for every subject. For the teaching the knowledge base, we set each number of input units and output units to three, and set the numbers of units in the middle layers to the numbers between 14 and 22 that have the lowest error for each subject.

3.2.3 Calculation of δ value

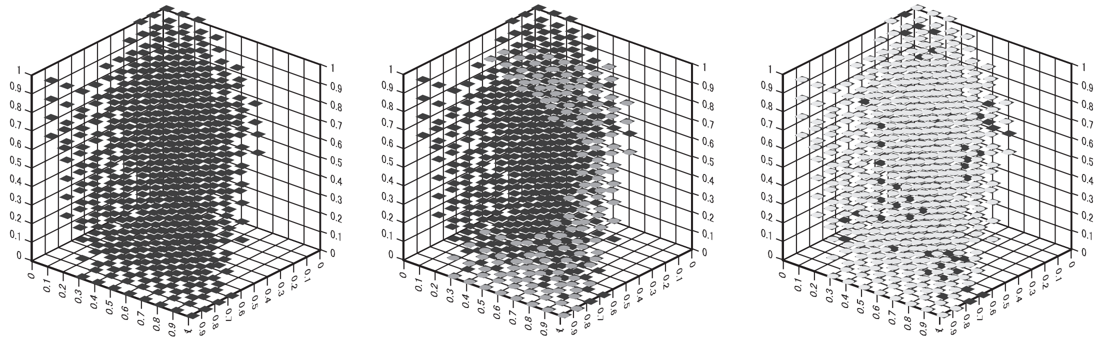
The output patterns of the three units are used as compression information. Thereby, compression information space can be considered to be 3D space with each side having the length of 1. We call this the 3D space compression information space. We divided the X, Y and Z-axis of this space at n intervals to investigate in detail the compression information in this space that satisfied each subject's preference. Each coordinates value of (i,j,k)$(0 \le i \le 1.0, 0 \le j \le 1.0, 0 \le k \le 1.0)$ is input into a knowledge base, and the square-sum of the difference between its output pattern and each value of a target preference result is calculated as δ value. We set the value of n to 10 (a unit of i, j, k are set to 0.1), and calculated the δ value. Here, each output pattern for each input pattern of compression information in the knowledge base is set to X_{out}, Y_{out} and Z_{out},and each unit value of preference results is set to X_a, Y_a, Z_a.We define the errors δ_{ijk} as $\delta_{ijk} = (X_a - X_{out_{ijk}})^2 + (Y_a - Y_{out_{ijk}})^2 + (Z_a - Z_{out_{ijk}})^2$.

3.2.4 Definition of δ space and visualization of δ space

We defined a four-dimensional space, which consisted of δ and each coordinate value of compression information space, as δ space. In fact, since human beings are unable to visualize four-dimensional space, we sliced the compression information space into eleven layers at the Z-axis. Then, we expressed the errors space as a figure, in which we set i as the X-axis, j as the Y-axis and δ_{ijk} as the Z-axis in each layer. An example of the error space for a certain subject, A, is shown in Figure 2. In this study, we considered that the compression information with an δ value of 0.01 or less was information that generated an optimum solution.

3.2.5 Exploring of optimum solutions

We extracted the compression information that generated optimum solutions from the visualized error space. The extracted



Optimum solutions

①Optimum solutions + DRs with each C.I. value in less than the 10th place of higher rank

②Optimum solutions + DRs with each C.I. value in less than the 50th place of higher rank

Legend

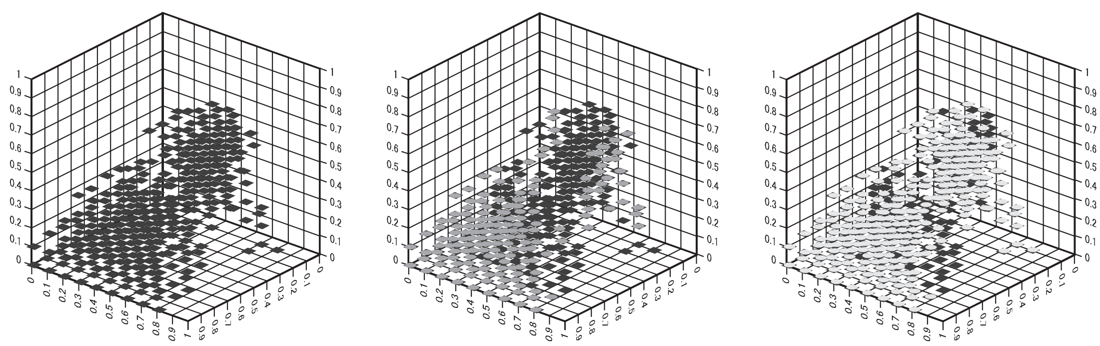
◆ Optimum solutions
● DRs with each C.I. value in less than the 10th place of higher rank
● DRs with each C.I. value in less than the 10th place of higher rank

Number of the preference samples	20
Number of DRs	84
Number of optimum solutions	655

Preferemce criteria

Item	Sound function	Recording speed of MD	Maximum output	MD changer	Price
Category	4	quadruple	31~51W	equipped	50,001~100,000yen

Figure. 3 Relation of optimum solutions and DRs of a certain subject included in type A



Optimum solutions

①Optimum solutions + DRs with each C.I. value in less than the 10th place of higher rank

②Optimum solutions + DRs with each C.I. value in less than the 50th place of higher rank

Legend

◆ Optimum solutions
● DRs with each C.I. value in less than the 10th place of higher rank
● DRs with each C.I. value in less than the 10th place of higher rank

Number of the preference samples	12
Number of DRs	64
Number of optimum solutions	355

Preferemce criteria

Item	Price	Audio devices	CD changer	Sound function	Recording speed of MD
Category	40,001～50,000 yen	3	3CDs	4	quadruple

Figure. 4 Relation of optimum solutions and DRs of a certain subject included in type B

compression information was input into the middle layer of the identity mapping model of functional information and the output patterns were calculated. We obtained these output patterns and designated them optimum solutions.

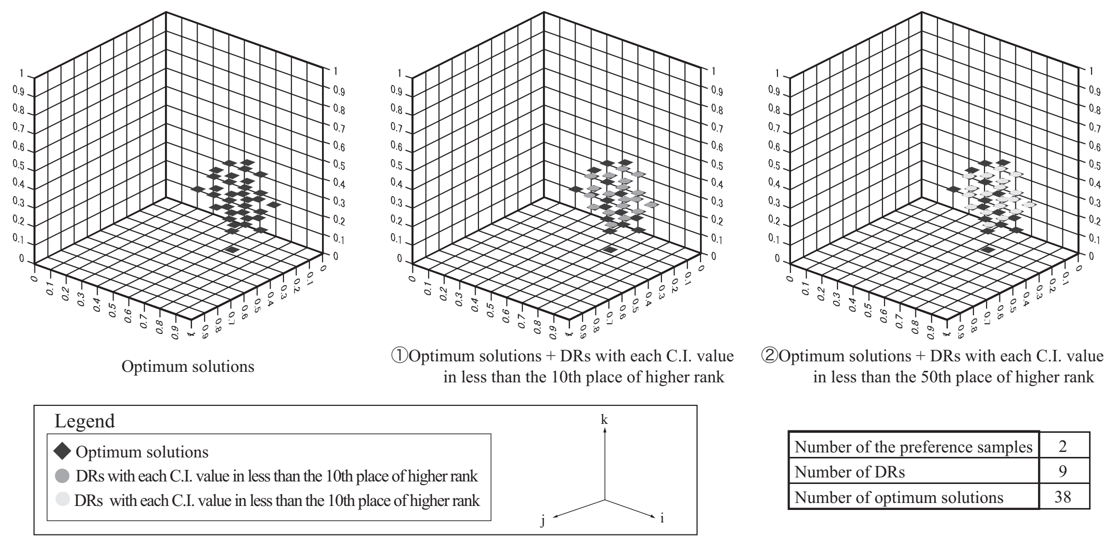
4. SIMULATION FOR COMPARISON

4.1 Comparison of Total Number of DRs, Optimum Solutions, and Preference Samples

In this study, we calculated DRs that derived from subjects' answers "I want to purchase" (a maximum of 100 items out of DRs having high C.I. values) and the optimum solutions that derived from subjects' answers that "I want to purchase" (δ are 0.01 or less) by applying rough set theory and an identity mapping model to the results of a survey of 40 subjects' preferences. Consequently, we found strong correlation among the total number of DRs, optimum solutions, and preference samples (see Table 2).

4.2 Comparison between DRs and Optimum Solutions

We compared DRs with optimum solutions, and investigated how DRs were included in optimum solutions. To compare them, the optimum solutions that were expressed the continuous amount were reapplied to the categories used in the rough set theory. When an optimum solution included



Number of the preference samples	2
Number of DRs	9
Number of optimum solutions	38

Preferemce criteria

Item	Recording speed of MD	Price	MD changer	CD changer	MDLP
Category	quadruple	40,001～50,000 yen	equipped	5or more CDs	equipped

Figure. 5 Relation of optimum solutions and DRs of a certain subject included in type C

one or more DRs, we considered that the DRs covered the optimum solutions. We compared the following two items (1) with optimum solutions, and we calculated the number of each item that covered optimum solutions.

(2) DRs with individual C.I. value in less than the 10th place of higher rank.
DRs with individual C.I. value in less than the 50th place of higher rank.

Here, the number of individual items that cover the optimum solutions varied with each subject. We considered that the number of individual items that cover the optimum solutions expressed subjects preferred features, as preference samples, DRs, or optimum solutions. Then, we observed the five items listed below, and we were able to classify 40 subjects' preference pattern into three clusters.

i) The total number of preference samples
ii) The total number of DRs
iii) The total number of optimum solutions
iv) The total number of the optimum solutions which DRs with each C.I. value in less than 10th place of higher rank cover
v) The total number of the optimum solutions which DRs with each C.I. value in less than 50th place of higher rank cover

We plotted the coordinates of compression information that derived optimum solutions in the compression information space, and we plotted the coordinates of compression information that generated optimum solutions that included DRs in the compression information space. The three types are listed below.

Type A), subjects with a nonlinear category as criteria of preference (See Figure 3)

- These subjects had many preference samples, a large total number of optimum solutions and large total number of DRs.
- "Price", "maximum outputs" and "CD changer" were often chosen as criteria of preference.
- The subjects tended to choose the categories that do not express "equipped" or "not equipped" as criteria of preference.

Type B), subjects with both linear criteria of preference and nonlinear criteria of preference (See Figure 4)

- "Price", "the number of the body", etc. were chosen as criteria of preference.
- Many DRs with individual C.I. value in less than 10th place of higher rank were included in optimum solutions.
- There were few differences in the number of optimum solutions which include between DRs with individual C.I. value in less than 10th place of higher rank and DRs with each C.I. value in less than 50th place of higher rank.

Type C), subjects with linear categories as criteria of preference (See Figure 5)

- The subjects had few numbers of preference samples, a small total number of optimum solutions, and small total number of DRs.
- The subjects tended to choose those categories that were expressed with two values like "equipped" or "not equipped" as criteria of preference.

5. COMPARISON EXPERIMENT AND CONSIDERATION

5.1 Comparison of Solutions Obtained from Each Method and Criteria of Preference

In this study, we judged whether the solutions obtained from

Table. 3 α value and β value

α value	β value
0.58	0.76

two methods would be appropriate by whether criteria of preference were included in them. Therefore, in order to investigate whether criteria of preference were included in DRs and the optimum solutions, we compared the following two items.

(1) Comparison between all DRs (a maximum of 50) and criteria of preference (five)

(2) Comparison between all optimum solutions (a maximum of 827) and criteria of preference (five)

When one or more criteria of preference were included in DRs or optimum solutions, we judged that DRs or optimum solutions covered criteria of preference, and we investigated about the following two items.

i) Average of the rate of DRs number including criteria of preference to total number of DRs $\rightarrow \alpha$ value

ii) Average of the rate of optimum solutions number including criteria of preference to total number of optimum solutions $\rightarrow \beta$ value

As a result, one or more criteria of preference was included in about 60 percent of DRs (See Table 3). Therefore, we clarified that we could extract preference attributes efficiently by using rough sets. Moreover, since criterion of preference was included in 70 percent or more of optimum solutions, the high reasoning capabilities of optimum solutions were also observed.

5.2 Comparison between the Categories which Many Include in the Solutions of Each Method and Criteria of Preference

We clarified whether the identity mapping model could extract a number of subjects' criteria of preference. However, since optimum solutions are obtained as a conjunction of many categories, it is unclear which categories are subjects' criteria of judgment to preference. Moreover, many optimum solutions were obtained. However, since a evaluation value of each optimum solution was only δ, which optimum solution should be used for a planning is unclear. Next, in this study, we considered the categories included in optimum solutions and DRs expressed strongly subjects' criteria of preference. We also considered that the maximum number of the categories that could actually be used for planning was ten. We conducted the following two comparison experiments for 18 subjects.

(1) Comparison of ten categories of higher rank that were included in all DRs and five criteria of preference

(2) Comparison of ten categories of higher rank that were included in all optimum solutions and five criteria of preference

Table 4 shows the average of DRs that were in agreement with the criteria of preference (D-P value), and optimum solutions of all subjects that were in agreement with the criteria of preference (O-P value).

Table. 4 D-P value and O-P value

D-P value	O-P value
2.88	1.12

5.3 The Relation of the Sample Used for Training and Categories Included in Many Optimum Solutions

In this study, we compared 10 categories that were included in optimum solutions and 10 categories of higher rank that were included in samples when the neural network trained. As a result, the categories that were included in optimum solutions and an average of 5.2 categories, which were included in many samples when the neural network trained were in agreement. Therefore, we conclude that optimum solutions are influenced by the deviation of the categories that were included in the sample used as the teacher signal for training.

6. CONCLUSION

Our findings in this study were the following

1) Strong correlation was found among the number of preference samples, DRs and optimum solutions.
2) We were able to divide subjects' preference patterns into three types.
3) Since many criteria of preference were found to be comparatively included in DRs and optimum solutions, we consider that rough sets and the identity mapping model are effective methods for the knowledge acquisition for planning.
4) The teacher signal must contain categories uniformly in every items. Otherwise, superfluous conformity will be produced.

The following subjects remain as future work.

1) We need to investigate in detail how DRs and optimum solutions are changed by subjects' criteria of preference and their answers.
2) We need to investigate in detail the subjects' sense of values by using the DRs obtained.

REFERENCES

[1] N.Mori, R.Takanashi. Design Support System Based on Reasoning with the Conception of Rough Set Theory. *Bulletin, Faculty of Arts, Tokyo Institute of Polytechnics*, pp.35-38, 1997.

[2] N.Zhong, J.Dong, S.Ohsuga. Rule Discovery in Medical Data by GDT-RS. *Journal of Japanese Society for Artificial Intelligence*, vol.15, no.5, pp.774-781, 2000.

[3] N.Takagi, K.Nakashima. Constructing and Minimizing Decision Rules Based on Heuristic Procedures. *Bulletin of International Rough Set Society*, vol.7, no.1/2, pp.23-27, 2003.

[4] Y.Enomoto, T.Harada. Analysis of Choice for Audio Products Using Annexation Reduct System. *Bulletin of Japanese Society for the Science of Design*, vol.49, no.5, pp.11-20, 2003.

[5] T.Harada, N.Mori. Development of Various Solution Using Identity Mapping Model. *Bulletin of Japanese Society for the Science of Design*, vol.41, no.1, pp.51-58, 1994.