Abstract—We applied a conceptually fuzzy set to a DVD recommendation system and evaluated it through experiments of DVD title recommendation. For experimental data, we used 12,000 movie titles and their content descriptions from Amazon.com. The experimental results show the proposed system could effectively recommend DVDs for a user.

I. INTRODUCTION

Recommendation systems use collaborative filtering. These systems use technology on the basis of the idea that things can be recommended to people on the basis of past choices of other people who share similar interests and tastes. For example, let us assume that the interests and tastes of user B resemble those of user A. Next, when user A is going to purchase an apple and an orange, if user B purchased liqueur in the past, the system would recommend that user a purchase liqueur. [1]. An example of this in practical use is merchandise recommendations on Amazon.com. [2]

There is also "Recommend.DRAGON2" from the DragonField Corporation and CbIR [3][4]. Recommend.DRAGON2 is a system that puts together image similarity recommendations, which recommend images that resemble other images, and choice record recommendations, which recommend choices on the basis of similar items chosen in the past.

Google AdWords, which the Google Corporation offers, is a tool examining what kind of keywords are used as search words with certain other keywords. For example, when a user searches "Christmas", Google AdWords shows associated words such as "present", "illuminations", etc. [5]

The recommendation system that we have developed uses a Conceptually Fuzzy Set (CFS). We expressed user's favor and noteworthy point conceptual by using CFS, and developed the system that recommended DVD filled with the user needs more accurately than a present recommendation system. In this article, we explain the CFS in Chapter 2. We then describe how we applied the technology to a recommendation system in Chapter 3 and give the experiment results in Chapter 4.

II. CONCEPTUALLY FUZZY SETS

A CFS is based on the use theory of meaning propounded by Wittgenstein for the expression of meaning of a concept. According to this theory, the meaning of a word can be expressed by another word. Thus, the meaning of a word can be expressed by other words associated with one another. Also, the expression of the meaning of a word by another word makes a closed circular system. [6]. According to his ideas, the meaning of a word can be expressed by another word. In other words, the meaning of any sentence can be expressed by another sentence. Moreover, the meaning of a statement expressed through another statement can be easily understood.

In CFS, the meaning of a word is expressed by the set of words and their activity values. General fuzzy sets express phenomena without a clear boundary. However, when applying them to various realistic problems, they are not situation-dependent. The meaning of a concept changes in situation-dependent phenomena, and cannot correspond to a fixed expression. This problem occurs when the generality of knowledge is not obtained because the degree the ambiguity of a fuzzy set is fixed, and the mechanism including background is not given. The essence of this problem is the expression of meaning. Conceptual fuzzy sets solve situation-dependent problems by a recall mechanism (i.e., combined fuzzy sets) and the general versatility of knowledge by making a closed circular system based on the use theory of meaning. [7] Using this idea, we have created two recommendation methods: A and B.

III. APPLICATION TO THE DVD RECOMMENDATION SYSTEM

A. An experiment object and recommendation algorithm

The data used in our recommendation system are only nouns extracted from the descriptions on amazon.com of DVDs as attribute data by Sen [8], as shown in figure 1, and are accompanied by a term frequency-inverse document frequency (TFIDF) score.

![Figure 1. Example of DVD data](image)

Operation history and DVD data are saved in a database, as shown in figure 2. In the current operation, the likes list and dislikes list that a user registers are constantly being saved. The user chooses DVDs he/she likes or does not like from DVD
titles displayed on a screen. This system revises the TFIDF score of the extracted words using this feedback. Of the words that are scored, the words with higher TFIDF scores are put into a recommendation engine.

TFIDF scores to a DVD title on the basis of input words and outputs a DVD list. Excluding DVD titles already registered, the system recommends other DVDs. The input is the words that have high scores. The output is DVD titles.

Figure 3 shows the how the database saves the input word group and information of whether DVDs recommended to a user are accepted or rejected.

**B. Method A**

In Method A, an atomic fuzzy set is generated by fuzzy clustering and a CFS is realized by the stack alignment. We perform fuzzy clustering for words in all DVD data; we use a cluster in which the words have a high degree of similarity and, to search for input words and degrees of similarity in each cluster, make an expansion language, as shown in figure 4.

After searching for the degrees of similarity in this expansion language and DVD data, the system recommends DVD titles that have a high degree of similarity. By deciding the meaning of words and characterizing them, the system can recommend DVDs on the basis of their content. This method depends on only words input at that time without depending on the user history, and it is used to collect histories to use in Method B.

**C. Method B**

The engine of Method B expresses relationships between the similar input words that appear together and recommends DVDs on the basis of them. As shown in figure 5, five words (red circles) are assumed to be input. At first, the system compares input words with input history saved in a database. Next, the engine adds the search histories that have a high degree of similarity and outputs a similar DVD title. Because the degree of similarity is measured by not the perfect matching but partial matches, a little recommendation of flexibility and the noise is possible.

As a result, a recommendation near the favor of the individual can be done by abstracting person's favor compared with the technique so far.

**IV. EXPERIMENT**

**A. Experimental data**

We tested our recommendation system to evaluate it. We acquired about 7,000 titles of popular DVDs from Amazon.com to use in the tests. We used Method A to collect 1,000 search histories in the laboratory. Next, we used Method B to recommend DVDs by using the 1,000 histories. We choose DVDs from the DVD titles that appear in the red frame in figure 6. Then the chosen DVD has its package / blurb displayed in the yellow frame. Next, the DVD is registered to the list of likes or
dislikes in the blue frame. Finally, the recommendation start button in the green frame is clicked.

Figure 6. Operation screen

B. Experiment 1

First of all, the recommendation that used Methods A and B was repeated several times.

B-1 Method A

In accordance with user preferences or interests, a DVD is recommended when a user repeats a search 4 or 5 times. It is thought that some characteristics catch the preferences or interests. Just like in the red frame in figure 7, an animation, such as "One Piece" or "Ghost In The Shell", is recommended by reference to another animation / children's story such as "Doraemon" or "A Christmas Carol". From this result, Method A captures content characteristics, and DVDs are recommended on the basis of features they share with other DVDs.

However, like the blue frames in figures 7 and 8, some words appear frequently, such as "ocean", regardless of what words are input. In other words, the recommendation results are biased.

Figure 7. Method A success example

Figure 8. Method A problems

This is because some clusters react easily and these clusters contain DVD titles with a high degree of similarity. To solve this problem, the selection of words used for the recommendation is reviewed, and this increases the number of clusters.

B-2 Method B

We search the engine of Method B by inputting our interests. As a result, it recommended DVD titles that we had expected and DVD titles that we had not expected but thought were interesting. We think it achieved our goal this time.

As a result of registering hard-boiled movies in the likes list, many similar movies were recommended.

Figure 9. Method B success example

From these results, it can be said that operation histories of other users who have similar interests were read, and DVDs were recommended by reference to the shared interests between users.

However, since we experimented only in our laboratory, there were some interests that are difficult to use when searching because of biased operation histories. For example, as shown in figure 10, though horror movies were registered in the likes list, very few horror movies were recommended. This is because few horror movies had been selected overall. We think the solution is to expand the number of users and to have a wider variety of search history contents.

C. Experiment 2

The recommendation results of cooperative filtering and a general cosine standard were compared with those of method B. The methods were evaluated using the following request as a
precondition: "An animated feature film that I do not know or one I might like."

**C-1 Cosine Standard**

The recommendation results were greatly influenced by the DVD data. Out of the ten recommended DVDs, only one was an animated feature film, and the expected results were not recommended.

**C-2 Collaborative Filtering**

In the recommendation results by using collaborative filtering, many animated feature films were recommended, probably because there was a user who likes animation a lot. However, no previously unknown animations were recommended because almost all DVDs were from the same series, such as Urusei Yatsura", "Doraemon", and "Crayon Shinchan".

**C-3 Method B**

In Method B, a lot of animations are recommended, and "Doraemon" was the most common. In addition, there were various recommendations such as "Finding Nemo", "Tales from Earthsea", and "Anpanman". When recommending "unknown, interesting animations" for the user, the proposed system produced excellent recommendation results.

The proposal system can be said that a flexible recommendation is possible than a usual system above.

**V. CONCLUSION AND FUTURE WORK**

The present study developed technology for recommending not only DVDs similar to those a user likes but also titles the user might not know about. Moreover, the improvement in searching for DVDs was confirmed by comparing the proposed method with cooperative filtering and a cosine standard. The proposed method produced many more DVD titles the user previously did not know about.

It can be said that the technique that uses the CFS system to express the feature of the item and person's idea is effective above. However, the volume of DVD data and search histories is too limited at present, and the system will need to be able to handle more information in the future. Moreover, a good quality recommendation system will be constructed by improving the word extraction method and the TFIDF scoring accuracy.

**REFERENCES**