World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:4, No:5, 2010

Multi-labeled Data Expressed by a Set of Labels

Tetsuya Furukawa and Masahiro Kuzunishi

Abstract—Collected data must be organized to be utilized efficiently, and hierarchical classification of data is efficient approach to organize data. When data is classified to multiple categories or annotated with a set of labels, users request multi-labeled data by giving a set of labels. There are several interpretations of the data expressed by a set of labels. This paper discusses which data is expressed by a set of labels by introducing orders for sets of labels and shows that there are four types of orders, which are characterized by whether the labels of expressed data includes every label of the given set of labels within the range of the set. Desirable properties of the orders, data is also expressed by the higher set of labels and different sets of labels express different data, are discussed for the orders.

Keywords—Classification Hierarchies, Multi-labeled Data, Multiple Classificaiton, Orders of Sets of Labels

I. Introduction

PROGRESS of information technologies and arrangement of network environments have been increasing available data including various kinds such as numerical data, texts, images, audio, etc. With the remarkable growth of data, it is becoming increasingly important to organize collected data properly. Hierarchical classification based on the content of data is one of the efficient methods to organize such data [2] [10] [11], which is used in the category searches in search engines, for example. Data is classified to categories or annotated with the labels of the categories.

Data is usually assumed to be classified to one category, which is called single-label classification [2] [14]. In *Newsgroups* data set, each news document is classified to only one category [12]. However, there is data which should be classified to multiple categories. For example, data on a comparison between manufacturing and financial industries should not be classified to either category *Manufacture* or *Finance* but to both in the classification for an industrial type. Such data is classified with multi-label classification, where data is classified to multiple categories [1] [8] [12]. In multi-label classification, the data on a comparison between manufacturing and financial industries is classified to both categories *Manufacture* and *Finance*, and labeled {*Manufacture, Finance*}.

Users or applications request data by giving labels. There are two kinds of "data identified by a label," the data with the same label as the given label and the data with a label whose concept is lower than or equal to the concept of the given label [7]. The data identified by label *Manufacture* is the data labeled *Manufacture* and the data with one of labels *Manufacture*, *Transportation*, *Automobile*, etc., respectively. In utilization of classified data, the latter is usually adopted, which this paper focuses on. When data is classified with single-label classification, the utilization of the data is rather straightforward. In multi-label classification, a set of labels can be used to identify a set of multi-labeled data because data have multiple labels. There are several kinds of "data identified by a set of labels."

T. Furukawa is Professor of Dept. of Economic Engineering, Kyushu University, Hakozaki 6–19–1, Higashi-ku, Fukuoka 812–8581 Japan (e-mail:furukawa@en.kyushu-u.ac.jp.)

M. Kuzunishi is Assistant Professor of Faculty of Business and Commerce, Aichi Gakuin University, Araike 12, Iwasaki, Nisshin, Aichi 470–0195 Japan (e-mail:kuzunisi@dpc.agu.ac.jp)

Example 1 Suppose a set of labels $\boldsymbol{L} = \{Manufacture, Finance\}$. The data identified by $\boldsymbol{L}$ is usually regarded as "the data related to nothing but manufacturing and financial industries" such as data labeled {*Automobile, Credit*}. On the other hand, there can be other sets of data identified by $\boldsymbol{L}$. When the data identified by $\boldsymbol{L}$ means "the data related to manufacturing and financial industries," it includes data labeled {*Automobile, Credit, Medicine*} where $Medicine$ is not related to $Manufacture$ or $Finance$. There are also such meanings that "the data related to only manufacturing industry or finance industry" and "the data related to manufacturing industry or finance industry," which include data labeled {*Automobile*} and {*Automobile, Medicine*}, with no label for $Finance$, respectively. □

Although there are several kinds of data identified by a set of labels, there is few discussions on the semantics shown in Example 1. Recent researches on classification allow multi-labeled data such as Web and texts [6] [11], whose purpose is automatic classification of data to multiple categories, and data is used through intersection or union of categories. In the utilization of multi-labeled data, methods to find the data matching given set of keywords are developed [3] [4], which rank data by frequency of keywords and their relationships so that users can find data satisfying their criteria. In those researches, the data identified by a given set of labels are such data as "the data related to all of the labels" or "the data related to any of the labels."

To utilize multi-labeled data precisely, there must be advanced usage based on the multiple labels. This paper introduces orders for sets of labels so that data is expressed by a set of labels if the label of the data is lower than or equal to the set of labels. Data is identified by a set of labels as the data expressed by the set of labels.

Usually a set of labels is interpreted as conjunction or disjunction of the elements, that is, the intersection or the union of the sets of data for the labels. These bring two types of orders for sets of labels. Other orders also exist, and those orders for sets of labels appear by systematic discussion. The purpose of this paper is to formalize the various possible orders.

There are two desirable properties of orders for sets of labels. The data identified by set of labels $\boldsymbol{L}_1$ should be also identified by set of labels $\boldsymbol{L}_2$ if $\boldsymbol{L}_1$ is lower than or equal to $\boldsymbol{L}_2$, and $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are generally expected to identify different data if $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are different from each other. These properties are discussed precisely.

This paper is organized as follows. Section 2 introduces orders for sets of labels. In Section 3, the data identified by sets of labels with the orders is discussed, and the orders are summarized to four types. Sections 4 and 5 discuss the properties of the orders to identify multi-labeled data. Section 6 concludes the paper.

II. Introducing Orders for Sets of Labels

Data is classified for each type of characteristic, which is called an attribute. For example, individual data is classied to the categories based on the industrial classification system, where the attribute is industry. While there is classification for multiple attributes [5] [11], this paper discusses one specific attribute for simplicity, and assumes that a classification hierarchy for the attribute is given and data is classified based on the hierarchy.

Let o be an object, an individual data, and L be a label which is used in classification of objects. Let $\overline{L}$ be the set of the objects expressed by L, and $\widetilde{o}$ be the label of o for the classification attribute. An object is classified to the lowest category (or categories in multi-label classification) corresponding to the object in a given classification hierarchy [1] [6] [8]. $\widetilde{o}$ is the label (or the set of labels) of the category (or the categories) to which o is classified. Objects may be classified to intermediate categories, which are not leaves in the hierarchy [6] [7] [13]. For example, if the hierarchies have the lower categories than manufacturing industry such as automobile industry, an object on the whole of manufacturing industry is not classified to the lower categories.

For labels L_1 and L_2, L_2 is higher than L_1 (L_1 is lower than L_2) if the category of L_2 is a higher concept of the category of L_1, denoted by $L_1 \prec L_2$. $L_1 \preceq L_2$ denotes that L_2 is higher than or equal to L_1. Thus $\prec$ is a partial order of labels given by a classification hierarchy. The membership of single-labeled objects to $\overline{L}$ is decided by the label of the objects as $\overline{L} = \{o \mid \widetilde{o} \preceq L\}$.

For multi-labeled objects, an order between a label and a set of labels have to be introduced to decide $\overline{L}$ because L is a label and $\widetilde{o}$ is a set of labels. Since a set of labels is usually interpreted as conjunction or disjunction of the elements, the orders for these interpretations are follows.

1) Conjunction: For a label L and a set of labels $\boldsymbol{L}$, $\boldsymbol{L}$ is lower than or equal to L if every label of $\boldsymbol{L}$ is lower than or equal to L, denoted by $\boldsymbol{L} \preceq_C L$.
2) Disjunction: For a label L and a set of labels $\boldsymbol{L}$, $\boldsymbol{L}$ is lower than or equal to L if some label of $\boldsymbol{L}$ is lower than or equal to L, denoted by $\boldsymbol{L} \preceq_D L$.

Example 2 Fig. 1 shows how sets of labels {*Automobile, Electronics*} and {*Automobile, Credit*} are lower than label *Manufacture*, where the dotted arcs from *Manufacture* to *Automobile* and *Electronics* express the order of the labels, *Manufacture* is higher than *Automobile* and *Electronics*. Since *Automobile* and *Electronics* are lower than *Manufacture*, {*Automobile, Electronics*} is lower than *Manufacture* for conjunction. {*Automobile, Credit*} is not because *Credit* is not lower than *Manufacture*. For disjunction, both {*Automobile, Electronics*} and {*Automobile, Credit*} are lower than *Manufacture* because they have lower labels of *Manufacture*. □

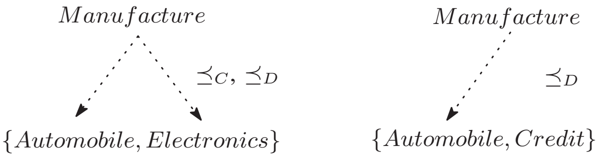

Fig. 1. Conjunction and Disjunction Interpretations of a Set of Labels

A label to express objects is extended to a set of labels. Let $\overline{\boldsymbol{L}}$ be the set of the objects expressed by a set of labels $\boldsymbol{L}$. Conjunction and disjunction interpretations of a set of labels for a label are extended to for a set of labels. Generally a set of labels $\boldsymbol{L}$ is interpreted as the intersection or the union of the sets of objects expressed by the labels of $\boldsymbol{L}$. Conjunction is extended at first. Let $\overline{\boldsymbol{L}}^{CI}$ and $\overline{\boldsymbol{L}}^{CU}$ be the intersection and the union of the sets of objects expressed by the labels in $\boldsymbol{L}$ for conjunction, which are the intersection and union interpretation of $\boldsymbol{L}$, respectively. They are formally expressed as

$$\overline{\boldsymbol{L}}^{CI} = \bigcap_{L \in \boldsymbol{L}} \{o \mid \widetilde{o} \preceq_C L\} \text{ and}$$
$$\overline{\boldsymbol{L}}^{CU} = \bigcup_{L \in \boldsymbol{L}} \{o \mid \widetilde{o} \preceq_C L\}.$$

Since the set of objects expressed by $\boldsymbol{L}$ is decided by the order of $\widetilde{o}$ and $\boldsymbol{L}$, orders for sets of labels have to be introduced. The orders corresponding to $\overline{\boldsymbol{L}}^{CI}$ and $\overline{\boldsymbol{L}}^{CU}$ are defined as follows.

Definition 1 For sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$,
$\boldsymbol{L}_1 \preceq_{CI} \boldsymbol{L}_2$ if $\forall L_2 \in \boldsymbol{L}_2, \forall L_1 \in \boldsymbol{L}_1, L_1 \preceq L_2$ and
$\boldsymbol{L}_1 \preceq_{CU} \boldsymbol{L}_2$ if $\exists L_2 \in \boldsymbol{L}_2, \forall L_1 \in \boldsymbol{L}_1, L_1 \preceq L_2$. □

The orders $\preceq_{CI}$ and $\preceq_{CU}$ exactly express $\overline{\boldsymbol{L}}^{CI}$ and $\overline{\boldsymbol{L}}^{CU}$, respectively.

Theorem 1 For a set of labels $\boldsymbol{L}$, $\overline{\boldsymbol{L}}^{CI} = \{o \mid \widetilde{o} \preceq_{CI} \boldsymbol{L}\}$ and $\overline{\boldsymbol{L}}^{CU} = \{o \mid \widetilde{o} \preceq_{CU} \boldsymbol{L}\}$. □

Proof: Since $\overline{\boldsymbol{L}}^{CI} = \bigcap_{L \in \boldsymbol{L}} \{o \mid \widetilde{o} \preceq_C L\}$, $\forall L \in \boldsymbol{L}, \widetilde{o} \preceq_C L$ for o in $\overline{\boldsymbol{L}}^{CI}$, that is, $\forall L \in \boldsymbol{L}, \forall L' \in \widetilde{o}, L' \preceq L$ for o in $\overline{\boldsymbol{L}}^{CI}$ by the definition of conjunction. Then $\overline{\boldsymbol{L}}^{CI}$ is expressed as $\{o \mid \forall L \in \boldsymbol{L}, \forall L' \in \widetilde{o}, L' \preceq L\}$, which is $\{o \mid \widetilde{o} \preceq_{CI} \boldsymbol{L}\}$ by Definition 1. In the same way, $\overline{\boldsymbol{L}}^{CU}$ is expressed as $\{o \mid \exists L \in \boldsymbol{L}, \forall L' \in \widetilde{o}, L' \preceq L\}$, which is $\{o \mid \widetilde{o} \preceq_{CU} \boldsymbol{L}\}$. *Q.E.D.*

In the same way as conjunction, disjunction is extended for a set of labels, and they are formally expressed as

$$\overline{\boldsymbol{L}}^{DI} = \bigcap_{L \in \boldsymbol{L}} \{o \mid \widetilde{o} \preceq_D L\} \text{ and}$$
$$\overline{\boldsymbol{L}}^{DU} = \bigcup_{L \in \boldsymbol{L}} \{o \mid \widetilde{o} \preceq_D L\}.$$

Definition 2 For sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$,
$\boldsymbol{L}_1 \preceq_{DI} \boldsymbol{L}_2$ if $\forall L_2 \in \boldsymbol{L}_2, \exists L_1 \in \boldsymbol{L}_1, L_1 \preceq L_2$ and

$\boldsymbol{L}_1 \preceq_{DU} \boldsymbol{L}_2$ if $\exists L_2 \in \boldsymbol{L}_2$, $\exists L_1 \in \boldsymbol{L}_1$, $L_1 \preceq L_2$. □

For a set of labels $\boldsymbol{L}$, the label of an object in $\overline{\boldsymbol{L}}^{DI}$ and $\overline{\boldsymbol{L}}^{DU}$ is lower than or equal to $\boldsymbol{L}$ according to $\preceq_{DI}$ and $\preceq_{DU}$, respectively.

Theorem 2 For a set of labels $\boldsymbol{L}$, $\overline{\boldsymbol{L}}^{DI} = \{o \mid \widetilde{o} \preceq_{DI} \boldsymbol{L}\}$ and $\overline{\boldsymbol{L}}^{DU} = \{o \mid \widetilde{o} \preceq_{DU} \boldsymbol{L}\}$. □

Proof: As the same as the proof of Theorem 1, $\overline{\boldsymbol{L}}^{DI}$ and $\overline{\boldsymbol{L}}^{DU}$ are expressed as $\{o \mid \forall L \in \boldsymbol{L}, \exists L' \in \widetilde{o}, L' \preceq L\}$ and $\{o \mid \exists L \in \boldsymbol{L}, \exists L' \in \widetilde{o}, L' \preceq L\}$, which are $\{o \mid \widetilde{o} \preceq_{DI} \boldsymbol{L}\}$ and $\{o \mid \widetilde{o} \preceq_{DU} \boldsymbol{L}\}$, respectively, by Definition 2. *Q.E.D.*

The orders for a label and a multi-labeled object were extended to the orders for sets of labels. There can be, on the other hand, the extension of orders for a set of labels and a single-labeled object.

There are two interpretations of a set of labels for single-labeled objects, intersection and union, which are formally expressed as $\bigcap_{L\in\boldsymbol{L}} \overline{L}$ and $\bigcup_{L\in\boldsymbol{L}} \overline{L}$, respectively.

Suppose intersection interpretation of $\boldsymbol{L}$. For a single-labeled object o and $L' = \widetilde{o}$, L' is lower than or equal to every label in $\boldsymbol{L}$, and $\overline{L'} \subseteq \bigcap_{L\in\boldsymbol{L}} \overline{L} = \bigcap_{L\in\boldsymbol{L}} \{o \mid \widetilde{o} \preceq L\}$. Thus a multi-labeled object o is expressed by $\boldsymbol{L}$ with conjunction if o is in $\bigcap_{L\in\boldsymbol{L}} \{o \mid \forall L' \in \widetilde{o}, L' \preceq L\}$. Let $\overline{\boldsymbol{L}}^{IC}$ be the set of objects expressed by $\boldsymbol{L}$ for this case, that is

$$\overline{\boldsymbol{L}}^{IC} = \bigcap_{L\in\boldsymbol{L}} \{o \mid \forall L' \in \widetilde{o}, L' \preceq L\}.$$

In the same way as $\overline{\boldsymbol{L}}^{IC}$, the sets of objects expressed by $\boldsymbol{L}$ for intersection interpretation of $\boldsymbol{L}$ with disjunction of multi-labeled objects, and for union interpretation of $\boldsymbol{L}$ with conjunction and disjunction of multi-labeled objects are defined as

$$\overline{\boldsymbol{L}}^{ID} = \bigcap_{L\in\boldsymbol{L}} \{o \mid \exists L' \in \widetilde{o}, L' \preceq L\},$$
$$\overline{\boldsymbol{L}}^{UC} = \bigcup_{L\in\boldsymbol{L}} \{o \mid \forall L' \in \widetilde{o}, L' \preceq L\}, \text{ and}$$
$$\overline{\boldsymbol{L}}^{UD} = \bigcup_{L\in\boldsymbol{L}} \{o \mid \exists L' \in \widetilde{o}, L' \preceq L\}.$$

Since the set of objects expressed by $\boldsymbol{L}$ consists of the objects whose label is lower than or equal to $\boldsymbol{L}$, the orders corresponding to $\overline{\boldsymbol{L}}^{IC}$, $\overline{\boldsymbol{L}}^{ID}$, $\overline{\boldsymbol{L}}^{UC}$, and $\overline{\boldsymbol{L}}^{UD}$ are introduced.

Definition 3 For sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$,
$\boldsymbol{L}_1 \preceq_{IC} \boldsymbol{L}_2$ if $\forall L_1 \in \boldsymbol{L}_1$, $\forall L_2 \in \boldsymbol{L}_2$, $L_1 \preceq L_2$,
$\boldsymbol{L}_1 \preceq_{ID} \boldsymbol{L}_2$ if $\exists L_1 \in \boldsymbol{L}_1$, $\forall L_2 \in \boldsymbol{L}_2$, $L_1 \preceq L_2$,
$\boldsymbol{L}_1 \preceq_{UC} \boldsymbol{L}_2$ if $\forall L_1 \in \boldsymbol{L}_1$, $\exists L_2 \in \boldsymbol{L}_2$, $L_1 \preceq L_2$, and
$\boldsymbol{L}_1 \preceq_{UD} \boldsymbol{L}_2$ if $\exists L_1 \in \boldsymbol{L}_1$, $\exists L_2 \in \boldsymbol{L}_2$, $L_1 \preceq L_2$. □

Theorem 3 For a set of labels $\boldsymbol{L}$, $\overline{\boldsymbol{L}}^{IC} = \{o \mid \widetilde{o} \preceq_{IC} \boldsymbol{L}\}$, $\overline{\boldsymbol{L}}^{ID} = \{o \mid \widetilde{o} \preceq_{ID} \boldsymbol{L}\}$, $\overline{\boldsymbol{L}}^{UC} = \{o \mid \widetilde{o} \preceq_{UC} \boldsymbol{L}\}$, and $\overline{\boldsymbol{L}}^{UD} = \{o \mid \widetilde{o} \preceq_{UD} \boldsymbol{L}\}$. □

Proof: $\overline{\boldsymbol{L}}^{IC}$ is defined as $\bigcap_{L\in\boldsymbol{L}} \{o \mid \forall L' \in \widetilde{o}, L' \preceq L\}$, which is $\{o \mid \forall L' \in \widetilde{o}, \forall L \in \boldsymbol{L}, L' \preceq L\}$. By the definition of $\preceq_{IC}$, $\overline{\boldsymbol{L}}^{IC} = \{o \mid \widetilde{o} \preceq_{IC} \boldsymbol{L}\}$. In the same way, $\overline{\boldsymbol{L}}^{UC}$ is $\{o \mid \forall L' \in \widetilde{o}, \exists L \in \boldsymbol{L}, L' \preceq L\}$, which is $\{o \mid \widetilde{o} \preceq_{UC} \boldsymbol{L}\}$ by the definition of $\preceq_{UC}$. The proofs of $\overline{\boldsymbol{L}}^{ID}$ and $\overline{\boldsymbol{L}}^{UD}$ are the same as the proofs of $\overline{\boldsymbol{L}}^{IC}$ and $\overline{\boldsymbol{L}}^{UC}$, respectively. *Q.E.D.*

III. The Objects Expressed by a Set of Labels

Section 2 introduced orders for sets of labels. This section shows what kinds of objects are expressed by a set of labels with those orders, and the orders are summarized to four types.

While an object o expressed by a set of labels $\boldsymbol{L}$ is decided by the order of $\boldsymbol{L}$ and $\widetilde{o}$, there may exist some labels in $\boldsymbol{L}$ and $\widetilde{o}$ which do nothing with the decision of the membership.

Example 3 Suppose $\boldsymbol{L}_1$ and $\widetilde{o_1}$ are {*Manufacture, Finance*} and {*Automobile, Credit, Medicine*}, respectively. o_1 is in $\overline{\boldsymbol{L}_1}^{DI}$ because there is a lower label in $\widetilde{o_1}$ for each label in $\boldsymbol{L}_1$. *Medicine* in $\widetilde{o_1}$ does nothing with this membership. Although there must be a label in $\widetilde{o_1}$ for each label of $\boldsymbol{L}_1$, $\widetilde{o_1}$ can include unrelated labels to $\boldsymbol{L}_1$. On the other hand, the labels of object o_2 labeled {*Automobile, Electronics*} in $\overline{\boldsymbol{L}_1}^{UC}$ are not lower than or equal to label *Finance* in $\boldsymbol{L}_1$. Object o_3 labeled {*Automobile, Medicine*} is in $\overline{\boldsymbol{L}_1}^{DU}$, where $Finance$ in $\boldsymbol{L}_1$ and $Medicine$ in $\widetilde{o_3}$ have no role for the membership of o_3 to $\overline{\boldsymbol{L}_1}^{DU}$. Fig. 2 illustrates these memberships. □

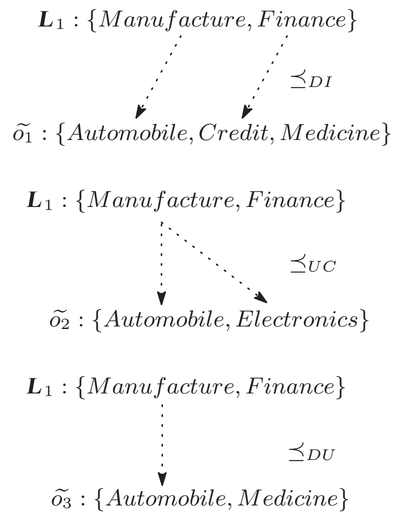

Fig. 2. Labels for Membership

For sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$, $\boldsymbol{L}_1 \preceq_{DI} \boldsymbol{L}_2$ requires that each label of $\boldsymbol{L}_2$ is lower than or equal to some label in $\boldsymbol{L}_1$, which is a restriction on the higher set $\boldsymbol{L}_2$. In the same way, $\boldsymbol{L}_1 \preceq_{UC} \boldsymbol{L}_2$ has the restriction on the lower set $\boldsymbol{L}_1$. There is no restriction in this meaning for $\boldsymbol{L}_1 \preceq_{DU} \boldsymbol{L}_2$, which is equivalent to $\boldsymbol{L}_1 \preceq_{UD} \boldsymbol{L}_2$. Thus $\preceq_{DI}$, $\preceq_{UC}$, and $\preceq_{DU}$ ($=\preceq_{UD}$) are renamed to $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RN}$, respectively.

$\boldsymbol{L}_1 \preceq_{RU} \boldsymbol{L}_2$ if $\forall L_2 \in \boldsymbol{L}_2, \exists L_1 \in \boldsymbol{L}_1, L_1 \preceq L_2$
$\boldsymbol{L}_1 \preceq_{RL} \boldsymbol{L}_2$ if $\forall L_1 \in \boldsymbol{L}_1, \exists L_2 \in \boldsymbol{L}_2, L_1 \preceq L_2$
$\boldsymbol{L}_1 \preceq_{RN} \boldsymbol{L}_2$ if $\exists L_1 \in \boldsymbol{L}_1, \exists L_2 \in \boldsymbol{L}_2, L_1 \preceq L_2$

Let $\overline{\boldsymbol{L}}^{RU}$, $\overline{\boldsymbol{L}}^{RL}$, and $\overline{\boldsymbol{L}}^{RN}$ be the sets of the objects expressed by a set of labels $\boldsymbol{L}$ with orders $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RN}$, respectively.

The rest of the orders are $\preceq_{ID}$, $\preceq_{IC}$, $\preceq_{CI}$, and $\preceq_{CU}$. For sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$, $\boldsymbol{L}_1 \preceq_{ID} \boldsymbol{L}_2$ and $\boldsymbol{L}_1 \preceq_{IC} \boldsymbol{L}_2$ when some and each label in $\boldsymbol{L}_1$ is lower than or equal to every label in $\boldsymbol{L}_2$, respectively. If $\boldsymbol{L}_2$ includes such labels L_{21} and L_{22} that $L_{21} \not\preceq L_{22}$ and $L_{22} \not\preceq L_{21}$, there does not exist such label L that $L \preceq L_{21}$ and $L \preceq L_{22}$. Since there is no label which is lower than or equal to every label in $\boldsymbol{L}_2$, any object is not expressed by $\boldsymbol{L}_2$ with $\preceq_{ID}$ or $\preceq_{IC}$. If $\boldsymbol{L}_2$ does not include such labels, $\boldsymbol{L}_2$ can be reduced to the lowest label in $\boldsymbol{L}_2$.

Example 4 Fig. 3 gives examples of memberships of objects o_1 to $\overline{\boldsymbol{L}}^{ID}$ (a) and o_2 to $\overline{\boldsymbol{L}}^{IC}$ (b), respectively. Label *Automobile* of $\widetilde{o_1}$ and each label *Automobile* and *Airplane* of $\widetilde{o_2}$ are lower than every label of $\boldsymbol{L}$. $\boldsymbol{L}$ can be reduced to $\boldsymbol{L}'$ which consists of the lowest label *Transportation*, because a label lower than or equal to *Transportation* is always lower than *Manufacture*. □

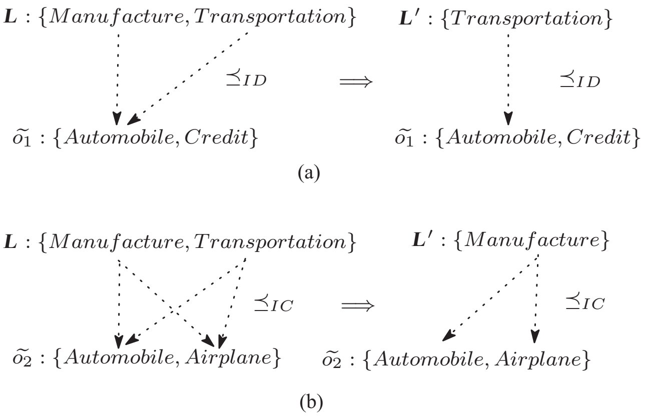

Fig. 3. Reduction of Labels

When a set of labels $\boldsymbol{L}$ is used to express objects with $\preceq_{ID}$ or $\preceq_{IC}$, $\boldsymbol{L}$ can be reduced to one label if $\overline{\boldsymbol{L}}^{ID} \neq \phi$ and $\overline{\boldsymbol{L}}^{IC} \neq \phi$. It is obvious that $\overline{\boldsymbol{L}}^{ID} = \overline{\boldsymbol{L}}^{RU}$ and $\overline{\boldsymbol{L}}^{IC} = \overline{\boldsymbol{L}}^{RL}$ when $|\boldsymbol{L}| = 1$. Since $\overline{\boldsymbol{L}}^{ID}$ and $\overline{\boldsymbol{L}}^{IC}$ are special cases of $\overline{\boldsymbol{L}}^{RU}$ and $\overline{\boldsymbol{L}}^{RL}$, respectively, $\preceq_{ID}$ and $\preceq_{IC}$ are excluded from our considerations. $\preceq_{CI}$ is also excluded because $\overline{\boldsymbol{L}}^{CI}$ is equal to $\overline{\boldsymbol{L}}^{IC}$.

For the last order $\preceq_{CU}$, $\overline{\{L_1\}}^{CU} \cap \overline{\{L_2\}}^{CU} = \phi$ if $L_1 \not\preceq L_2$ and $L_2 \not\preceq L_1$, and $\overline{\{L_1\}}^{CU} \subseteq \overline{\{L_2\}}^{CU}$ if $L_1 \preceq L_2$, for L_1 and L_2 in $\boldsymbol{L}$. Thus $\overline{\boldsymbol{L}}^{CU} = \bigcup_{L \in \boldsymbol{L}} \overline{\{L\}}^{CU}$ is the direct union of the objects expressed by the labels in $\boldsymbol{L}$ which have no higher label in $\boldsymbol{L}$. Since the labels in $\boldsymbol{L}$ are treated individually and $\overline{\{L\}}^{CU} = \overline{\{L\}}^{RL}$, $\preceq_{CU}$ is also excluded from the discussion.

The orders proposed in Section 2 are reduced to $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RN}$. There may be other orders defined as that a set of labels $\boldsymbol{L}_1$ are lower than or equal to a set of labels $\boldsymbol{L}_2$ if $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_2$ and $\boldsymbol{L}_1 \preceq_y \boldsymbol{L}_2$ ($x, y \in \{CI, CU, DI, DU, IC, ID, UC, UD\}$). The orders except the order defined with $x = DI$ and $y = UC$ are either $\preceq_x$ or $\preceq_y$. For example, the order defined with $x = CI$ and $y = CU$ is $\preceq_{CI}$.

Since $\preceq_{DI}$ and $\preceq_{UC}$ are $\preceq_{RU}$ and $\preceq_{RL}$, respectively, the order where $x = DI$ and $y = UC$ has restrictions of $\preceq_{RU}$ and $\preceq_{RL}$. Such order is denoted by $\preceq_{RB}$, where $\preceq_{RB}$ restricts both of higher and lower sets of labels. Let $\overline{\boldsymbol{L}}^{RB}$ be the set of objects expressed by a set of labels $\boldsymbol{L}$ with order $\preceq_{RB}$. Since $\overline{\boldsymbol{L}}^{RB}$ is expressed as $\overline{\boldsymbol{L}}^{RB} = \{o \mid \tilde{o} \preceq_{RB} \boldsymbol{L}\} = \{o \mid \tilde{o} \preceq_{RU} \boldsymbol{L}, \tilde{o} \preceq_{RL} \boldsymbol{L}\}$, $\preceq_{RB}$ is defined as follows.

For sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$, $\boldsymbol{L}_1 \preceq_{RB} \boldsymbol{L}_2$ if every label of $\boldsymbol{L}_2$ is higher than or equal to some labels of $\boldsymbol{L}_1$ and every label of $\boldsymbol{L}_1$ is lower than or equal to some labels of $\boldsymbol{L}_2$.

$$\boldsymbol{L}_1 \preceq_{RB} \boldsymbol{L}_2 \text{ if } \forall L_2 \in \boldsymbol{L}_2, \exists L_1 \in \boldsymbol{L}_1, L_1 \preceq L_2 \text{ and } \forall L_1 \in \boldsymbol{L}_1, \exists L_2 \in \boldsymbol{L}_2, L_1 \preceq L_2$$

The objects expressed by a set of labels $\boldsymbol{L}$ are $\overline{\boldsymbol{L}}^{RN}$, $\overline{\boldsymbol{L}}^{RU}$, $\overline{\boldsymbol{L}}^{RL}$, and $\overline{\boldsymbol{L}}^{RB}$. $\overline{\boldsymbol{L}}^{RN}$ and $\overline{\boldsymbol{L}}^{RL}$ are the union of the objects expressed by the labels of $\boldsymbol{L}$, and $\overline{\boldsymbol{L}}^{RU}$ and $\overline{\boldsymbol{L}}^{RB}$ are the intersection of the objects expressed by the labels of $\boldsymbol{L}$. $\overline{\boldsymbol{L}}^{RN}$ and $\overline{\boldsymbol{L}}^{RU}$ include objects with labels which are not related to $\boldsymbol{L}$, and $\overline{\boldsymbol{L}}^{RL}$ and $\overline{\boldsymbol{L}}^{RB}$ do not. In the other words, the labels of the objects in $\overline{\boldsymbol{L}}^{RL}$ and $\overline{\boldsymbol{L}}^{RB}$ are within the range of $\boldsymbol{L}$. These discussions are summarized in Fig. 4.

		Range	
		No	Yes
Interpretation	Union	RN	RL
	Intersection	RU	RB

Fig. 4. Interpretation and Rage of Sets of Labels

Example 5 For set of labels $\boldsymbol{L} = \{Manufacture, Finance\}$, $\overline{\boldsymbol{L}}^{RN}$ and $\overline{\boldsymbol{L}}^{RL}$ are the union of the objects expressed by the labels of $\boldsymbol{L}$, which include objects labeled {*Automobile*}, {*Automobile, Credit*}, {*Automobile, Credit, Medicine*}, etc. for $\overline{\boldsymbol{L}}^{RN}$ and {*Automobile*}, {*Automobile, Credit*}, etc. for $\overline{\boldsymbol{L}}^{RL}$. $\overline{\boldsymbol{L}}^{RU}$ and $\overline{\boldsymbol{L}}^{RB}$ are the intersection, which include the objects labeled {*Automobile, Credit*}, {*Automobile, Credit, Medicine*}, etc. for $\overline{\boldsymbol{L}}^{RU}$ and {*Automobile, Credit*}, etc. for $\overline{\boldsymbol{L}}^{RB}$. While objects of $\overline{\boldsymbol{L}}^{RN}$ and $\overline{\boldsymbol{L}}^{RU}$ may include label *Medicine* which is not related to *Manufacture* or *Finance*, the labels of objects of $\overline{\boldsymbol{L}}^{RL}$ and $\overline{\boldsymbol{L}}^{RB}$ are within the range of *Manufacture* and *Finance*. □

IV. Soundness of Orders

In Section 3, the orders for sets of labels were summarized to four types by discussing the objects expressed by sets of labels. This section shows a desirable property of the orders for sets of labels to express multi-labeled objects.

In single-label classification, the order of labels is defined by the order of categories in a classification hierarchy. A label L_1 is lower than a label L_2 when the category for L_1 is lower than the category for L_2. Since a classification hierarchy expresses concepts in a hierarchical order, the order of labels agrees with the order of concepts. Thus it is naturally accepted that an object in $\overline{L_1}$ is in $\overline{L_2}$ if L_1 is lower than or equal to L_2. In multi-label classification, the concept of a set of labels is not clear. If an order for sets of labels agrees with the order for the concepts of sets of labels as the same as single-label classification, an object in $\overline{\boldsymbol{L}_1}$ is expected to be in $\overline{\boldsymbol{L}_2}$ for such sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ that $\boldsymbol{L}_1$ is lower than or equal to $\boldsymbol{L}_2$.

Definition 4 An order $\preceq_x$ for sets of labels is sound if $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_2$ is equivalent to $\overline{\boldsymbol{L}_1}^x \subseteq \overline{\boldsymbol{L}_2}^x$ for any sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$. □

Order $\preceq_{RN}$ is not sound. Suppose sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ such that $\boldsymbol{L}_1 \preceq_{RN} \boldsymbol{L}_2$.There may exist a label L_1 in $\boldsymbol{L}_1$ which is not lower than or equal to any label of $\boldsymbol{L}_2$. While an object which has a label lower than or equal to L_1 is in $\overline{\boldsymbol{L}_1}^{RN}$, the object may not be in $\overline{\boldsymbol{L}_2}^{RN}$ because the object may not have a label which is lower than or equal to a label of $\boldsymbol{L}_2$. Thus there can exist such objects that are in $\overline{\boldsymbol{L}_1}^{RN}$ but not in $\overline{\boldsymbol{L}_2}^{RN}$.

Example 6 Let sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ be {*Manufacture, Credit*} and $\{Finance\}$, respectively. Since *Credit* in $\boldsymbol{L}_1$ is lower than *Finance* in $\boldsymbol{L}_2$, $\boldsymbol{L}_1 \preceq_{RN} \boldsymbol{L}_2$. Although object o labeled $\{Automobile\}$ is in $\overline{\boldsymbol{L}_1}^{RN}$ because *Automobile* is lower than *Manufacture*, o is not in $\overline{\boldsymbol{L}_2}^{RN}$ because *Automobile* is not lower than or equal to *Finance*. Fig. 5 illustrates the orders between $\boldsymbol{L}_1$, $\boldsymbol{L}_2$, and $\widetilde{o}$. □

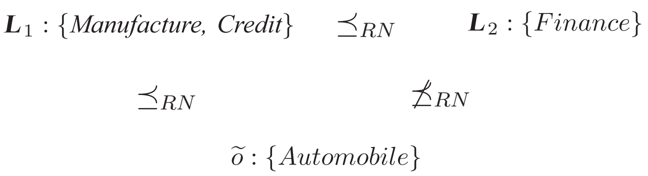

Fig. 5. Membership for $\preceq_{RN}$

The transitivity of orders is a necessary and sufficient condition for the soundness of orders.

Lemma 1 An order is sound if and only if the order is transitive. □

Proof: Suppose an order $\preceq_x$ is transitive. For a set of labels $\boldsymbol{L}_1$, an object o is in $\overline{\boldsymbol{L}_1}^{x}$ if $\widetilde{o} \preceq_x \boldsymbol{L}_1$. o is also in such $\overline{\boldsymbol{L}_2}^{x}$ that $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_2$ because $\widetilde{o} \preceq_x \boldsymbol{L}_2$ by the transitivity of $\widetilde{o} \preceq_x \boldsymbol{L}_1$ and $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_2$. Since every object in $\overline{\boldsymbol{L}_1}^{x}$ is also in $\overline{\boldsymbol{L}_2}^{x}$, $\overline{\boldsymbol{L}_1}^{x} \subseteq \overline{\boldsymbol{L}_2}^{x}$. If $\overline{\boldsymbol{L}_1}^{x} \subseteq \overline{\boldsymbol{L}_2}^{x}$, object o in $\overline{\boldsymbol{L}_1}^{x}$ is also in $\overline{\boldsymbol{L}_2}^{x}$. $\widetilde{o} \preceq_x \boldsymbol{L}_2$, and $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_2$ when $\widetilde{o} = \boldsymbol{L}_1$. Thus $\preceq_x$ is sound if $\preceq_x$ is transitive.

For any sets of labels $\boldsymbol{L}_1$, $\boldsymbol{L}_2$, and $\boldsymbol{L}_3$ such that $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_2$ and $\boldsymbol{L}_2 \preceq_x \boldsymbol{L}_3$, $\overline{\boldsymbol{L}_1}^{x} \subseteq \overline{\boldsymbol{L}_2}^{x}$ and $\overline{\boldsymbol{L}_2}^{x} \subseteq \overline{\boldsymbol{L}_3}^{x}$ if $\preceq_x$ is sound. Since $\overline{\boldsymbol{L}_1}^{x} \subseteq \overline{\boldsymbol{L}_2}^{x} \subseteq \overline{\boldsymbol{L}_3}^{x}$, an object o in $\overline{\boldsymbol{L}_1}^{x}$ is in $\overline{\boldsymbol{L}_3}^{x}$. $o \preceq_x \boldsymbol{L}_3$, and $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_3$ when $\widetilde{o} = \boldsymbol{L}_1$. Thus $\preceq_x$ is transitive if $\preceq_x$ is sound. *Q.E.D.*

While $\preceq_{RN}$ is not transitive as shown in Example 6, where $\widetilde{o} \preceq_{RN} \boldsymbol{L}_1$ and $\boldsymbol{L}_1 \preceq_{RN} \boldsymbol{L}_2$ but $\widetilde{o} \not\preceq_{RN} \boldsymbol{L}_2$, $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RB}$ are transitive.

Lemma 2 Orders $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RB}$ are transitive. □

Proof: For sets of labels $\boldsymbol{L}_1$, $\boldsymbol{L}_2$, and $\boldsymbol{L}_3$ such that $\boldsymbol{L}_1 \preceq_{RU} \boldsymbol{L}_2$ and $\boldsymbol{L}_2 \preceq_{RU} \boldsymbol{L}_3$, $\forall L_3 \in \boldsymbol{L}_3, \exists L_1 \in \boldsymbol{L}_1, L_1 \preceq L_3$ because $\forall L_2 \in \boldsymbol{L}_2$, $\exists L_1 \in \boldsymbol{L}_1$, $L_1 \preceq L_2$ and $\forall L_3 \in \boldsymbol{L}_3$, $\exists L_2 \in \boldsymbol{L}_2$, $L_2 \preceq L_3$. Thus $\boldsymbol{L}_1 \preceq_{RU} \boldsymbol{L}_3$, and $\preceq_{RU}$ is transitive. The proofs for $\preceq_{RL}$ and $\preceq_{RB}$ are as the same as for $\preceq_{RU}$. *Q.E.D.*

Order $\preceq_{RU}$, $\preceq_{RL}$ and $\preceq_{RB}$ are transitive, and soundness of them is proved.

Theorem 4 Orders $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RB}$ are sound. □

Proof: $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RB}$ are transitive by Lemma 2 and sound by Lemma 1. *Q.E.D.*

V. Proper Orders for Sets of Labels

Another desirable property of the orders for sets of labels is discussed in this section. Set of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are generally expected to express different objects when $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ is different from each other.

Definition 5 An order is proper if $\overline{\boldsymbol{L}_1} \neq \overline{\boldsymbol{L}_2}$ for any different sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$. □

Let L be a label in $\boldsymbol{L}_1 - \boldsymbol{L}_2$ for sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$. The objects expressed by $\boldsymbol{L}_1$ are generally different from the objects expressed by $\boldsymbol{L}_2$ because of L. If there is a label in $\boldsymbol{L}_1 \cap \boldsymbol{L}_2$ which is lower than or equal to L, there does not exist such object that is in $\overline{\boldsymbol{L}_1}^{RU}$ but not in $\overline{\boldsymbol{L}_2}^{RU}$ because L is in $\boldsymbol{L}_1$.

Example 7 Let sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ be $\{Manufacture, Automobile\}$ and $\{Automobile\}$, respectively. Although $\boldsymbol{L}_1 - \boldsymbol{L}_2$ is $\{Manufacture\}$, there does not exist such object in $\overline{\boldsymbol{L}_1}^{RU}$ that is not in $\overline{\boldsymbol{L}_2}^{RU}$ because *Automobile* in $\boldsymbol{L}_1 \cap \boldsymbol{L}_2$ is lower than *Manufacture*. □

The resulted orders in Section 3 are proper if sets of labels are limited to that there is no labels L_i and L_j of $L_i \preceq L_j$ in a set of labels. Such set of labels are called exclusive. However, there are sets of labels which are not exclusive but should be considered. For example, the label of an object on the share of automobile industy in manufacturing industry must be $\{Manufacture, Automobile\}$, which is not exclusive.

There may exist such sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ ($\boldsymbol{L}_1 \neq \boldsymbol{L}_2$) that $\boldsymbol{L}_1 \preceq_{RU} \boldsymbol{L}_2$ and $\boldsymbol{L}_2 \preceq_{RU} \boldsymbol{L}_1$, denoted by $\boldsymbol{L}_1 \approx_{RU} \boldsymbol{L}_2$, if $\boldsymbol{L}_1$ or $\boldsymbol{L}_2$ is not exclusive.

Example 8 Let $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ be $\{Transportation, Finance\}$ and $\{Manufacture, Transportation, Finance\}$, respectively. $\boldsymbol{L}_1 \approx_{RU} \boldsymbol{L}_2$ because $\boldsymbol{L}_1 \preceq_{RU} \boldsymbol{L}_2$ and $\boldsymbol{L}_2 \preceq_{RU} \boldsymbol{L}_1$ as shown in Fig. 6. □

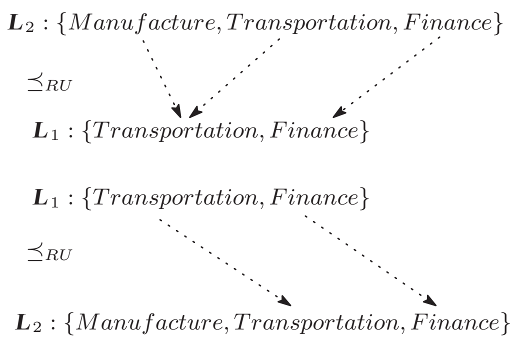

Fig. 6. Example for $\approx_{RU}$

Suppose $\boldsymbol{L}_1 \preceq_{RU} \boldsymbol{L}_2$ for sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$, and let L_2 and L_2' be such labels that $L_2 \in \boldsymbol{L}_2$, $L_2' \notin \boldsymbol{L}_2$, and

$L_2 \preceq L'_2$. $\boldsymbol{L}_1 \preceq_{RU} \boldsymbol{L}_2 \cup \{L'_2\}$ $(= \boldsymbol{L}'_2)$ because $L_1 \preceq L'_2$ for any label L_1 in $\boldsymbol{L}_1$ such that $L_1 \preceq L_2$ (Fig. 7). Thus $\overline{\boldsymbol{L}_2}^{RU}$ and $\overline{\boldsymbol{L}'_2}^{RU}$ is the same set of objects.

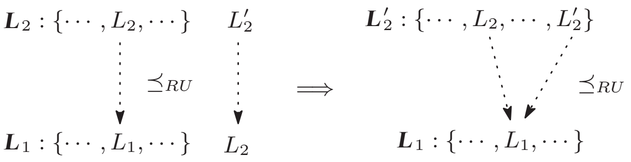

Fig. 7. Redundant Labels of $\preceq_{RU}$

Generally, a set of labels $\boldsymbol{L}$ can be reduced to the subset of $\boldsymbol{L}$ consisting of the labels which are not higher than any other labels of $\boldsymbol{L}$ for $\preceq_{RU}$. Such subset is defined as the lower bound of $\boldsymbol{L}$, formally expressed as

$$l(\boldsymbol{L}) = \{L \mid L \in \boldsymbol{L}, \forall L' \in \boldsymbol{L} \ (L' \neq L), L' \not\preceq L\}.$$

Lemma 3 For a set of labels $\boldsymbol{L}$, $\boldsymbol{L} \approx_{RU} l(\boldsymbol{L})$. □

Proof: Since there exists such L' in $l(\boldsymbol{L})$ that $L' \preceq L$ for each label L in $\boldsymbol{L}$, $\forall L \in \boldsymbol{L}$, $\exists L' \in l(\boldsymbol{L})$, $L' \preceq L$, which is the definition of $l(\boldsymbol{L}) \preceq_{RU} \boldsymbol{L}$. Since $l(\boldsymbol{L})$ is a subset of $\boldsymbol{L}$, $\forall L \in l(\boldsymbol{L})$, $\exists L' \in \boldsymbol{L}$, $L = L'$, and $\boldsymbol{L} \preceq_{RU} l(\boldsymbol{L})$. Thus $\boldsymbol{L} \approx_{RU} l(\boldsymbol{L})$. *Q.E.D.*

The objects expressed by a set of labels $\boldsymbol{L}$ with $\preceq_{RU}$ is the same objects expressed by the lower bound of $\boldsymbol{L}$.

Theorem 5 For a set of labels $\boldsymbol{L}$, $\overline{\boldsymbol{L}}^{RU} = \overline{l(\boldsymbol{L})}^{RU}$. □

Proof: Each object o in $\overline{\boldsymbol{L}}^{RU}$ is also in $\overline{l(\boldsymbol{L})}^{RU}$ because $\widetilde{o} \preceq_{RU} \boldsymbol{L} \approx_{RU} l(\boldsymbol{L})$ by Lemma 3, and $\overline{\boldsymbol{L}}^{RU} \subseteq \overline{l(\boldsymbol{L})}^{RU}$. $\overline{l(\boldsymbol{L})}^{RU} \subseteq \overline{\boldsymbol{L}}^{RU}$ because each object o in $\overline{l(\boldsymbol{L})}^{RU}$ is in $\overline{\boldsymbol{L}}^{RU}$ by $\widetilde{o} \preceq_{RU} l(\boldsymbol{L}) \approx_{RU} \boldsymbol{L}$. Thus $\overline{\boldsymbol{L}}^{RU} = \overline{l(\boldsymbol{L})}^{RU}$. *Q.E.D.*

For sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ $(\boldsymbol{L}_1 \neq \boldsymbol{L}_2)$, $\overline{\boldsymbol{L}_1}^{RU} = \overline{\boldsymbol{L}_2}^{RU}$ if $l(\boldsymbol{L}_1) = l(\boldsymbol{L}_2)$ by Theorem 5, which shows that $\preceq_{RU}$ is not proper.

In the same way as the lower bound of a set of labels, the upper bound of a set of labels $\boldsymbol{L}$ is introduced for $\overline{\boldsymbol{L}}^{RL}$. The upper bound of $\boldsymbol{L}$ is the subset of $\boldsymbol{L}$ consisting of the labels which is not lower than any labels of $\boldsymbol{L}$, formally expressed as

$$u(\boldsymbol{L}) = \{L \mid L \in \boldsymbol{L}, \forall L' \in \boldsymbol{L} \ (L' \neq L), L \not\preceq L'\}.$$

Since the same theorems for $\overline{\boldsymbol{L}}^{RL}$ and $\overline{u(\boldsymbol{L})}^{RL}$ and for $\overline{\boldsymbol{L}}^{RN}$ and $\overline{u(\boldsymbol{L})}^{RN}$ as Theorem 5 can be proved, orders $\preceq_{RL}$ and $\preceq_{RN}$ are not proper.

$\preceq_{RB}$ is not proper either because there exists such sets of labels $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ that $\boldsymbol{L}_1 \approx_{RB} \boldsymbol{L}_2$.

Example 9 Let $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ be $\{Manufacture, Automobile\}$ and $\{Manufacture, Transportation, Automobile\}$, respectively. $\boldsymbol{L}_1 \approx_{RB} \boldsymbol{L}_2$ because $\boldsymbol{L}_1 \preceq_{RB} \boldsymbol{L}_2$ and $\boldsymbol{L}_2 \preceq_{RB} \boldsymbol{L}_1$, which is shown in Fig. 8. □

Let $ul(\boldsymbol{L})$ be $u(\boldsymbol{L}) \cup l(\boldsymbol{L})$. $\boldsymbol{L}_2$ in Example 9 can be reduced to $\boldsymbol{L}_1$ for $\preceq_{RB}$, which is $ul(\boldsymbol{L}_2)$.

Theorem 6 For a set of labels $\boldsymbol{L}$, $\overline{\boldsymbol{L}}^{RB} = \overline{ul(\boldsymbol{L})}^{RB}$. □

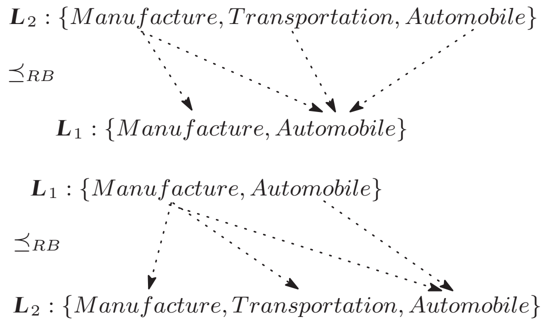

Fig. 8. Example for $\approx_{RB}$

Proof: Since $ul(\boldsymbol{L})$ includes $u(\boldsymbol{L})$, there exists such L' in $ul(\boldsymbol{L})$ that $L \preceq L'$ for each label L in $\boldsymbol{L}$, and $\forall L \in \boldsymbol{L}$, $\exists L' \in ul(\boldsymbol{L})$, $L \preceq L'$. $\forall L \in ul(\boldsymbol{L})$, $\exists L' \in \boldsymbol{L}$, $L' \preceq L$ because $ul(\boldsymbol{L})$ is a subset of $\boldsymbol{L}$. Consequently, $\boldsymbol{L} \preceq_{RB} ul(\boldsymbol{L})$, and every object in $\overline{\boldsymbol{L}}^{RB}$ is also in $\overline{ul(\boldsymbol{L})}^{RB}$, that is, $\overline{\boldsymbol{L}}^{RB} \subseteq \overline{ul(\boldsymbol{L})}^{RB}$. In the same way, since $ul(\boldsymbol{L})$ includes $l(\boldsymbol{L})$, there exists such L' in $ul(\boldsymbol{L})$ that $L' \preceq L$ for each label L in $\boldsymbol{L}$, and $\forall L \in \boldsymbol{L}$, $\exists L' \in ul(\boldsymbol{L})$, $L' \preceq L$. $\forall L \in ul(\boldsymbol{L})$, $\exists L' \in \boldsymbol{L}$, $L \preceq L'$ because $ul(\boldsymbol{L})$ is a subset of $\boldsymbol{L}$. Consequently, $ul(\boldsymbol{L}) \preceq_{RB} \boldsymbol{L}$, and every object in $\overline{ul(\boldsymbol{L})}^{RB}$ is also in $\overline{\boldsymbol{L}}^{RB}$, that is, $\overline{ul(\boldsymbol{L})}^{RB} \subseteq \overline{\boldsymbol{L}}^{RB}$. *Q.E.D.*

The objects expressed by a set of labels $\boldsymbol{L}$ with $\preceq_{RU}$, $\preceq_{RL}$ and $\preceq_{RN}$, and $\preceq_{RB}$, are the same objects expressed by $l(\boldsymbol{L})$, $u(\boldsymbol{L})$, and $ul(\boldsymbol{L})$, respectively. Thus a set of labels is reduced to the upper or the lower bound of the sets of labels when the set is not exclusive.

VI. Conclusion

This paper showed that the objects expressed by a set of labels $\boldsymbol{L}$ are $\overline{\boldsymbol{L}}^{RN}$, $\overline{\boldsymbol{L}}^{RU}$, $\overline{\boldsymbol{L}}^{RL}$, and $\overline{\boldsymbol{L}}^{RB}$. The difference of them is due to the interpretation of $\widetilde{o}$ and $\boldsymbol{L}$, whether $\widetilde{o}$ is within the range of $\boldsymbol{L}$ and whether $\boldsymbol{L}$ express intersection or union, which were formally discussed by introducing orders for sets of labels.

There were two desirable properties of orders. One is that $\preceq_{RU}$, $\preceq_{RL}$, and $\preceq_{RB}$ are sound, that is, $\boldsymbol{L}_1 \preceq_x \boldsymbol{L}_2$ is equivalent to $\boldsymbol{L}_1 \subseteq \boldsymbol{L}_2$. Since the objects expressed by a set of labels $\boldsymbol{L}_1$ is also expressed by a set of labels $\boldsymbol{L}_2$ if $\boldsymbol{L}_1$ is lower than or equal to $\boldsymbol{L}_2$ with these orders, the orders can be used for the concepts of sets of labels.

If sets of labels are exclusive, every order is proper, where different sets of labels express different sets of objects. Since labels of objects are generally not exclusive, sets of labels should not be limited to be exclusive. In utilization of such objects, sets of labels are reduced to the lower and upper bounds of the sets for $\preceq_{RU}$, and $\preceq_{RL}$ and $\preceq_{RN}$, respectively, and sets of labels are reduced to the union of the lower and upper bounds of the sets for $\preceq_{RB}$.

This paper gave framework to utilize multi-labeled objects with multiple labels, which can use for advanced application. In the fields such as semantic web and knowledge management, we often face multi-label classification and utilization of multi-labeled data [1] [8] [12]. The results of this paper can be applied to such fields.

REFERENCES

[1] G. Adami, P. Avesani, and D. Sona, "Bootstrapping for Hierarchical Document Classification," *Proc. Int'l Conf. on on Information and Knowledge Management* (CIKM'03), pp. 295–302, 2003.
[2] E. Bertino, J. Fan, E. Ferrari, M. Hachi, and A. Elamagarmid, "A Hierarchical Access Control Model for Video Database Systems," *ACM Transactions on Information Systems*, Vol.21, No.2, pp. 151–191, 2003.
[3] K. Chakrabarti, V. Ganti, J. Han, and D. Xin, "Rankig Objects by Exploiting, Relationships: Computing Top-K over Aggregation," *Proc. ACM SIGMOD Int'l Conf. on Management of Data*, pp. 371–382, 2006.
[4] S. Chuang and L. Chien, "Taxonomy Generation for Text Segments:A Practical Web-Based Approach," *ACM Transactions on Information Systems*, Vol.23, No.4, pp. 363–396, 2005.
[5] W. Dakka, P. G. Ipeirotis, and K. R. Wood, "Automatic Construction of Multifaceted Browsing Interfaces," *Proc. Int'l Conf. on Information and Knowledge Management* (CIKM'05), pp. 768–775, 2005.
[6] S. Dumais and H. Chen, "Hierarchical Classification of Web Content," *Proc. ACM Int'l Conf. on Research and Development in Information Retrieval*, pp. 256–263, 2000.
[7] T. Furukawa and M. Kuzunishi, "Hierarchical Classification of Heterogeneous Data," *Proc. IASTED Int'l Conf. on Databases and Applications* (DBA2005), pp. 252–257, 2005.
[8] N. Ghamrawi and A. McMallum, "Collecticve Multi-Label Classification," *Proc. Int'l Conf. on Information and Knowledge Management* (CIKM'05), pp. 195–200, 2005.
[9] D. Koller and M. Sahami, "Hierarchically Classifying Documents Using Very Few Words," *Proc. Int'l Conf. on Machine Learning*, pp. 170–178, 1997.
[10] M. Kuzunishi and T. Furukawa, "Representation for Multiple Classified Data," *Proc. IASTED Int'l Conf. on Databases and Applications* (DBA2006), pp. 135–142, 2006.
[11] A. Sun and E. Lim, "Hierarchical Text Classification and Evaluation," *Proc. IEEE Int'l Conf. on Data Mining* (ICDM2001), pp. 521–528, 2001.
[12] K. Toutanova, F. Chen, K. Popat, and T. Hofmann, "Text Classification in a Hierarchical Mixture Model for Small Training Sets," *Proc. Int'l Conf. on Information and Knowledge Management* (CIKM'01), pp. 105–112, 2001.
[13] K. Wang, S. Zhou, and Y. He, "Hierarchical Classification of Real Life Documents," *Proc. SIAM Int'l Conf. on Data Mining*, pp. 1–16, 2001.
[14] K. Wang, S. Zhou, and S. C. Liew, "Building Hierarchical Classifiers Using Class Proximity," *Proc. Int'l Conf. on Very Large Data Bases* (VLDB'99), pp. 363–374, 1999.