Granular Computing Based on a Generalized Approximation Space

Jian-Min Ma[1], Wen-Xiu Zhang[1], Wei-Zhi Wu[2], and Tong-Jun Li[1,2]

[1] Institute for Information and System Sciences, Faculty of Science, Xi'an Jiaotong University, Xi'an, Shaan'xi 710049, P.R. China
majm@mail.xjtu.edu.cn, wxzhang@mail.xjtu.edu.cn
[2] Mathematics, Physics and Information College, Zhejiang Ocean University, Zhoushan, Zhejiang, 316004, P.R. China
wuwz8681@sina.com, litj@zjou.net.cn

Abstract. A family of overlapping granules can be formed by granulating a finite universe under a binary relation in a set-theoretic setting. In this paper, we granulate a universe by a binary relation and obtain a granular universe. And then we define two kinds of operators between these two universes, study properties of them. By combining these two kinds of operators, we get two pairs of approximation operators. It is proved that one kind of combination operators is just the approximation operators under a generalized approximation space defined according to Pawlak's rough set theory.

Keywords: Generalized approximation space, $L-$lower approximation operator, $H-$upper approximation operator, Similarity relation.

1 Introduction

Granular computing is a label of theories, methodologies, techniques, and tools that makes use of granules, i.e., groups, classes, or clusters of a universe, in the process of problem solving [10, 14, 16]. Since Pawlak introduced the theory of rough sets [7,8], it has made granular computing popular . Hobbs [2] introduced the concepts of granularity in 1985. Later the concept "granular computing" was suggested by Zadeh [15, 16] for the first time in 1996. The basic ideas of information granulation have been explored in many fields, such as rough sets, fuzzy sets, cluster analysis, database, machine learning, data -mining, and so on. There is a renewed and fast growing interest in the study of granular computing [3, 4, 10, 13].

As a concrete theory of granular computing, rough set model enables us to precisely define and analyze many notions of granular computing. The results provide an in-depth understanding of granular computing. Many models of granular computing have been proposed and studied [16, 11]. However, there are many fundamental issues in granular computing, such as granulation of the universe, description of granules, relationships between granules, and computing with granules.

J.T. Yao et al. (Eds.): RSKT 2007, LNAI 4481, pp. 93–100, 2007.

Yao [12] proposed a concrete model of granular computing based on a simple granulation structure, namely, a partition of a universe. Results from rough sets, quotient space theory, belief functions, and power algebra are reformulated, reinterpreted, and combined for granular computing. For the universe and the coarse-grained universe induced by an equivalence relation, two basic operation called zooming-out and zooming-in operations are introduced. And Computations in these universes can be connected through the two operations.

Because the equivalence relation in [12] is too strong to be obtained in general, we only consider a reflexive relation on a universe which is easy to obtain usually. Then a covering model can be obtained by granulating a finite set of a universe based on the reflexive relation [6]. And we cited definitions of zooming-out and zooming-in operations in [12] and discussed the covering model of granular computing [6]. However, relationships between subsets of a coarse-grained universe would not hold in the universe. Furthermore, although rough set approximations of a classical subset of a universe in a generalized approximation space [17] can be obtained by a combination of these operations, the duality may not hold.

In this paper, we first granulate a finite set of a universe into a family of overlapping granules based on a general binary relation. We introduce two kinds of operators between a universe and the granulated universe, and study their properties. Then we combine them to two pairs of approximation operators, which are used to study connections between computations in the two universes. It is also proved that approximation representations of a generalized approximation space can be obtained by combining them, and the duality always holds for the different combinations.

This paper is organized as follows. Section 2 introduces two kinds of operators between a universe and a granulated universe, and studies their properties. Section 3 shows new operations formed by different combining the two operations, investigates their properties, and discusses connections between computations in the two universes. Finally, Section 4 concludes the paper.

2 Preliminaries

Let U be a finite and nonempty set called a universe, and $r \subseteq U \times U$ a binary relation on the universe U. For any $x \in U$, the set $r(x) = \{y \in U;\ (x, y) \in r\}$ is called the successor neighborhood of x. The relation r is referred to as serial if for any $x \in U$, there exists $y \in U$ such that $y \in r(x)$. r is referred to as reflexive if for all $x \in U,\ x \in r(x)$; r is referred to as symmetric if for all $x, y \in U,\ x \in r(y)$ implies $y \in r(x)$; r is referred to as transitive if for all $x, y, z \in U,\ x \in r(y)$ and $y \in r(z)$ implies $x \in r(z)$; r is referred to as Euclidean if for all $x, y, z \in U,\ y \in r(x)$ and $z \in r(x)$ implies $z \in r(y)$ [9,17]. Furthermore, r is referred to as a similarity relation on U if it is reflexive and transitive; r is referred to as a tolerance relation on U if it is reflexive and symmetric. For any binary relation $r \subseteq U \times U$, the pair (U, r) is referred to as a generalized approximation space.

For a generalized approximation space (U, r), the family of all successor neighborhood, denoted by $\mathcal{A} = \{r(x);\ x \in U\}$, is commonly knows as the granulated set. If r is reflexive, it forms a covering of U, namely, a family of overlapping subsets whose union is U. For any $x \in U$, the successor neighborhood $r(x)$ is considered as a whole granule instead of many individuals [12]. It is a subset of U and an element of $\mathcal{A}$. We use $|r(x)|$ to denote the whole granule $r(x)$, and call $A = \{|r(x)|;\ x \in U\}$ a granulated universe. We denote by 2^U the power set of the universe U, and by ^{c} the set complement operator.

Definition 1. *Let (U, r) be a generalized approximation space. For any $B \subseteq A$, the mapping $f : 2^A \to 2^U$ is given by*

$$f(B) = \{x \in U;\ |r(x)| \in B\}.$$

Then we can get the following properties: for any $B, C \subseteq A$,
(1) $f(\emptyset) = \emptyset$;
(1) $f(A) = U$;
(2) $f(B \cup C) = f(B) \cup f(C)$;
(3) $f(B \cap C) = f(B) \cap f(C)$;
(4) $f(B^c) = f(B)^c$;
(5) $B \subseteq C \Leftrightarrow f(B) \subseteq f(C)$.

Definition 2. *Let (U, r) be a generalized approximation space and $X \subseteq U$. A pair (L, H) of mappings $L, H : 2^U \to 2^A$ is defined as follows:*

$$L(X) = \{|r(x)|;\ r(x) \subseteq X\},$$
$$H(X) = \{|r(x)|;\ r(x) \cap X \neq \emptyset\}.$$

They are called $L-$lower and $H-$upper approximation of X, respectively.

By Definition 2 we can easily get that for a generalized approximation space (U, r), $X, Y \subseteq U$,
(LH1) $H(\emptyset) = \emptyset$;
(LH2) $L(U) = A$;
(LH3) $L(X^c) = H(X)^c$, $H(X^c) = L(X)^c$;
(LH4) $L(X \cap Y) = L(X) \cap L(Y)$, $H(X \cap Y) \subseteq H(X) \cap H(Y)$;
(LH5) $L(X \cup Y) \supseteq L(X) \cup L(Y)$, $H(X \cup Y) = H(X) \cup H(Y)$;
(LH6) $X \subseteq Y \Rightarrow L(X) \subseteq L(Y)$, $H(X) \subseteq H(Y)$;
(LH7) Let $B_n(X) = H(X) - L(X)$, then $B_n(X^c) = B_n(X)$.
If r is serial, we have $L(X) \subseteq H(X)$, $L(\emptyset) = \emptyset$ and $H(U) = A$.

Remark 1. In fact, there exists $X \subseteq U$ and $X \neq \emptyset$ such that $L(X) = \emptyset$; and there is $X \neq U$ such that $H(X) = A$.

In general, the following formulas may not hold:

$$H(X \cap Y) = H(X) \cap H(Y),$$
$$L(X \cup Y) = L(X) \cup L(Y).$$

Example 1. Suppose $U = \{x_1, x_2, x_3, x_4, x_5\}$, and $r \subseteq U \times U$ be a binary relation on U satisfying: $r(x_1) = \{x_3, x_4\}$, $r(x_2) = \{x_1, x_2, x_4\}$, $r(x_3) = \{x_3\}$, $r(x_4) = \{x_4\}$, $r(x_5) = \{x_2, x_5\}$.

(1) Take $X = \{x_3, x_4\}$ and $Y = \{x_2, x_3\}$. Then $X \cap Y = \{x_3\}$ and $H(X \cap Y) = \{|r(x_1)|, |r(x_3)|\}$, but $H(X) \cap H(Y) = \{|r(x_1)|, |r(x_2)|, |r(x_3)|, |r(x_4)|\} \cap \{|r(x_1)|, |r(x_2)|, |r(x_3)|, |r(x_5)|\} = \{|r(x_1)|, |r(x_2)|, |r(x_3)|\}$. Hence $H(X \cap Y) \subset H(X) \cap H(Y)$.

(2) Take $X = \{x_1, x_2\}$ and $Y = \{x_3\}$. Then $X \cup Y = \{x_1, x_2, x_3\}$ and $L(X \cup Y) = \{|r(x_3)|, |r(x_5)|\}$, however $L(X) \cup L(Y) = \emptyset \cup \{|r(x_3)|\} = \{|r(x_3)|\}$. Hence $L(X) \cup L(Y) \subset L(X \cup Y)$.

Proposition 1. *Let* (U, r) *be a generalized approximation space and* $X, Y \subseteq U$. *Note that* $\underline{Z}(X, Y) = \{|r(x)|;\ r(x) \subseteq X \cup Y, |r(x)| \in B_n(X) \cap B_n(Y)\}$. *Then*

$$L(X \cup Y) = L(X) \cup L(Y) \cup \underline{Z}(X, Y).$$

Proof. It is easy to see that $L(X) \cup L(Y) \subseteq L(X \cup Y)$. Then $|r(x)| \in L(X \cup Y) - L(X) \cup L(Y)$ if and only if $r(x) \subseteq X \cup Y$, $r(x) \nsubseteq X$ and $r(x) \nsubseteq Y$. Then $|r(x)| \in L(X \cup Y) - L(X) \cup L(Y)$ if and only if $r(x) \subseteq X \cup Y$, and $|r(x)| \in B_n(X) \cap B_n(Y)$. That is $L(X \cup Y) - L(X) \cup L(Y) = \underline{Z}(X, Y)$. Therefore $L(X \cup Y) = L(X) \cup L(Y) \cup \underline{Z}(X, Y)$.

Proposition 2. *Let* (U, r) *be a generalized approximation space and* $X, Y \subseteq U$. *Note that* $\overline{Z}(X, Y) = \{|r(x)|;\ r(x) \cap (X \cap Y) = \emptyset, |r(x)| \in B_n(X) \cap B_n(Y)\}$. *Then*

$$H(X \cap Y) = H(X) \cap H(Y) - \overline{Z}(X, Y).$$

Proof. By (LH5) we can get that $H(X \cap Y) \subseteq H(X) \cap H(Y)$. Then $|r(x)| \in H(X) \cap H(Y) - H(X \cap Y)$ if and only if $r(x) \cap X \neq \emptyset, r(x) \cap Y \neq \emptyset$ and $r(x) \cap (X \cap Y) = \emptyset$. Then $|r(x)| \in H(X) \cap H(Y) - H(X \cap Y)$ if and only if $|r(x)| \in B_n(X) \cap B_n(Y)$, and $r(x) \cap (X \cap Y) = \emptyset$. Therefore $H(X) \cap H(Y) - H(X \cap Y) = \overline{Z}(X, Y)$. Thus $H(X \cap Y) = H(X) \cap H(Y) - \overline{Z}(X, Y)$.

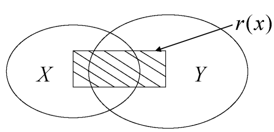

Fig. 1. $\underline{Z}(X, Y)$

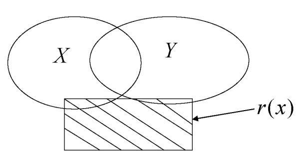

Fig. 2. $\overline{Z}(X, Y)$

3 Rough Approximation Representations on *U* and *A*

Pawlak's classical rough set theory shows that lower and upper approximations of a classical set are also subsets of the same universe. From Definition 1 and

Definition 2 we can combine these operators f and (L, H), and get some new operators on a same universe.

By Definitions 1 and 2, one can easily obtain lower and upper approximations of a classical subset of the generalized approximation space by performing a combination of (L, H) and f as follows:

$$fL(X) = f(\{|r(x)|;\ r(x) \subseteq X\}) = \{x;\ r(x) \subseteq X\},$$
$$fH(X) = f(\{|r(x)|;\ r(x) \cap X \neq \emptyset\}) = \{x;\ r(x) \cap X \neq \emptyset\}.$$

Then $fL, fH : 2^U \to 2^U$ are called $fL-$lower and $fH-$upper approximation operators, respectively.

Since we have studied properties of f and (L, H), we can easily get the following properties for any $X, Y \subseteq U$:

(fLH1) $fH(\emptyset) = \emptyset$;
(fLH2) $fL(U) = U$;
(fLH3) $fL(X^c) = fH(X)^c$, $fH(X^c) = fL(X)^c$;
(fLH4) $fL(X \cap Y) = fL(X) \cap fL(Y)$,
$fH(X \cap Y) \subseteq fH(X) \cap fH(Y)$;
(fLH5) $fL(X \cup Y) \supseteq fL(X) \cup fL(Y)$,
$fH(X \cup Y) = fH(X) \cup fH(Y)$;
(fLH6) $X \subseteq Y \Rightarrow fL(X) \subseteq fL(Y)$, $fH(X) \subseteq fH(Y)$;
(fLH7) Let $B_1(X) = fH(X) - fL(X)$, then $B_1(X^c) = B_1(X)$.

Note that (fLH7) means simply that if we cannot decide when an object is in X, we obviously cannot decide whether it is in X^c either. (fLH3) shows that fL and fH are the dual approximation operators.

If r is reflexive, then $fL(X) \subseteq X \subseteq f(H(X)$, $fL(\emptyset) = \emptyset$,$fH(U) = U$.

Proposition 3. *Let (U, r) be a generalized approximation space. If r is reflexive and Euclidean, for any $X \subseteq U$, we have*
(1) $fH(X) = fL(fH(X))$;
(2) $fL(X) = fH(fL(X))$.

Proof. Since r is reflexive, we have $fL(fH(X)) \subseteq fH(X)$. Take $x \in fH(X)$. Then $r(x) \cap X \neq \emptyset$. For any $y \in r(x)$, we have $r(x) \subseteq r(y)$ because r is Euclidean. Hence $r(y) \cap X \neq \emptyset$ and $y \in fH(X)$. By the arbitrariness of y we can get $r(x) \subseteq fH(X)$ which leads to $x \in fL(fH(X))$. Therefore $fL(fH(X)) = fH(X)$. By (1) and (fLH3) we can easily get (2).

According to Propositions 1 and 2 we can get:

Proposition 4. *Let (U, r) be a generalized approximation space. Then for any $X, Y \subseteq U$, we have*
(1) $fL(X \cup Y) = fL(X) \cup fL(Y) \cup \underline{Z}_1(X, Y)$, where
$\underline{Z}_1(X, Y) = \{x \in U;\ r(x) \subseteq X \cup Y,\ r(x) \subseteq B_1(X) \cap B_1(Y)\}$.
(2) $fH(X \cap Y) = fH(X) \cap fH(Y) - \overline{Z}_1(X, Y)$, where
$\overline{Z}_1(X, Y) = \{x \in U;\ r(x) \cap (X \cap Y) = \emptyset,\ r(x) \subseteq B_1(X) \cap B_1(Y)\}$.

For a subset $B \subseteq A$ we can obtain a subset $f(B) \subseteq U$, and then obtain a pair of subsets $Lf(B)$ and $Hf(B)$ as follows:

$$Lf(B) = \{|r(x)|;\ r(x) \subseteq f(B)\} = \{|r(x)|;\ r(x) \subseteq \{y; |r(y)| \in B\}\},$$

$$Hf(B)) = \{|r(x)|;\ r(x) \cap f(B) \neq \emptyset\} = \{|r(x)|;\ \exists y \in r(x), |r(y)| \in B\}.$$

Then $Lf, Hf : 2^A \to 2^A$ are called $Lf-$lower and $Hf-$upper approximation operators, respectively. And these two approximation operators give out the approximation representations of B in the granulated universe A.

So for any $B, C \subseteq A$ we have the following properties:

(LHf1) $Hf(\emptyset) = \emptyset$;
(LHf2) $Lf(A) = A$;
(LHf3) $Lf(B^c) = (Hf(B))^c$, $Hf(B^c) = (Lf(B))^c$;
(LHf4) $Lf(B \cap C) = Lf(B) \cap Lf(C)$,
$Hf(B \cap C) \subseteq Hf(B) \cap Hf(C)$;
(LHf5) $Lf(B \cup C) \supseteq Lf(B) \cup Lf(C)$,
$Hf(B \cup C) = Hf(B) \cup Hf(C)$;
(LHf6) $B \subseteq C \Rightarrow Lf(B) \subseteq Lf(C)$, $Hf(B) \subseteq Hf(C)$;
(LHf7) Let $B_2(B) = Hf(B) - Lf(B)$, then $B_2(B^c) = B_2(B)$.

If r is serial, $Lf(\emptyset) = \emptyset$ and $Hf(A) = A$; if r is reflexive, we have $Lf(B) \subseteq B \subseteq Hf(B)$.

For any $X \subseteq U$ and $B \subseteq A$, by the different combinations of f and (L, H) we can get

$$fL(f(B)) = f(Lf(B)),$$

$$Lf(L(X)) = L(fL(X)).$$

Therefore, we can easily get

Lemma 1. *Let (U, r) be a generalized approximation space. If r is a similarity relation, for any $X \subseteq U$ we have*
(1) $L(fL(X)) = L(X)$;
(2) $H(fH(X)) = H(X)$.

Proof. Since r is reflexive, $L(fL(X)) \subseteq L(X)$. Conversely, take $|r(x)| \in L(X)$. Then $r(x) \subseteq X$. Since r is transitive, for any $y \in r(x)$ we have $r(y) \subseteq r(x)$. Therefore $|r(y)| \in L(X)$ and $y \in fL(X)$. By the arbitrariness of y we can get $r(x) \subseteq fL(X)$. Hence $|r(x)| \in L(fL(X))$, and $L(X) \subseteq L(fL(X))$. From which we get $L(fL(X)) = L(X)$. By (LHf3) we can prove (2).

However, the following formulae may not hold:

$$fL(f(B)) = f(B),$$

$$fH(f(B)) = f(B).$$

Example 2. Suppose $U = \{x_1, x_2, x_3, x_4, x_5\}$, and r is a similarity relation on U satisfying: $r(x_1) = \{x_1, x_3, x_4\}$, $r(x_2) = \{x_2\}$, $r(x_3) = \{x_3, x_4\}$, $r(x_4) = \{x_4\}$, $r(x_5) = \{x_2, x_5\}$. Then $A = \{|r(x_1)|, |r(x_2)|, |r(x_3)|, |r(x_4)|, |r(x_5)|\}$.

(1) Take $B = \{|r(x_1)|, |r(x_3)|\}$, then $f(B) = \{x_1, x_3\}$, and $L(f(B)) = L(\{x_1, x_3\}) = \emptyset$. But $fL(f(B)) = f(\emptyset) = \emptyset \neq f(B)$.
(2) Take $B = \{|r(x_4)|, |r(x_5)|\}$, then $f(B) = \{x_4, x_5\}$, and $H(f(B)) = \{|r(x_1)|, |r(x_3)|, |r(x_4)|, |r(x_5)|\}$. However $fH(f(B)) = f(\{|r(x_1)|, |r(x_3)|, |r(x_4)|, |r(x_5)|\}) = \{x_1, x_3, x_4, x_5\} \neq f(B)$.

Proposition 5. *Let (U, r) be a generalized approximation space. If r is a similarity relation, then for any $B \subseteq A$ we have*
(1) $Lf(Lf(B)) = Lf(B)$;
(2) $Hf(Hf(B)) = Hf(B)$.

In addition, according to Propositions 1, 2 and properties of f we can get:

Proposition 6. *Let (U, r) be a generalized approximation space. If r is reflexive, then for any $B, C \subseteq A$,*
(1) $Lf(B \cup C) = Lf(B) \cup Lf(C) \cup \underline{Z}_2(B, C)$; *where*
$\underline{Z}_2(B, C) = \{|r(x)|;\ r(x) \subseteq f(B \cup C), r(x) \subseteq B_2(B) \cap B_2(C)\}$.
(2) $Hf(B \cap C) = Hf(B) \cap Hf(C) - \overline{Z}_2(B, C)$; *where*
$\overline{Z}_2(B, C) = \{|r(x)|;\ r(x) \cap f(B \cap C) = \emptyset, r(x) \subseteq B_2(B) \cap B_2(C)\}$.

Literatures [7, 17] define lower and upper approximation operators for a generalized approximation space (U, R) with R being a binary relation on U as follows:

$$\underline{R}(X) = \{x \in U; R_s(x) \subseteq X\},$$
$$\overline{R}(X) = \{x \in U; R_s(x) \cap X \neq \emptyset\},$$

where $R_s(x)$ denotes the successor neighborhood of x. Obviously, for a generalized approximation space (U, r), $fL(X) = \underline{r}(X)$ and $fH(X) = \overline{r}(X)$. Since we have studied properties of operators f and (L, H), we can easily get properties of $(\underline{r}(X), \overline{r}(X))$.

4 Conclusion

Granular computing is a way of thinking that relies in our ability to perceive the real world under various grain sizes, to abstract and consider only those things that serve our present interest, and to switch among different granularities. In this paper, two kinds of operators have been introduced between a universe and a granulated universe based on a generalized binary relation. Connections between the elements of a universe and the elements of a granulated universe, as well as connections between computations in the two universes are investigated by two pairs of combination operators.

Acknowledgement

This work was supported by a grant from the National Natural Science Foundation of China (No. 60673096), the National 973 Program of China (No. 2002CB 312200) and the Scientific Research Project of the Education Department of Zhejiang Province in China (No. 20061126).

References

1. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems, 17 (1990) 191-208
2. Hobbs, J.R.: Granularity. In: Proc. of IJCAI, Los Angeles, (1985) 432-435
3. Lin, T.Y.: Granular computing on binary relations I: data mining and neighborhood systems, II: rough set representations and belief functions. In: Rough Sets in Knowledge Discovery 1. Polkowski, L., Skowron, A(Eds.) Physica-Verlag, Heidelberg (1998) 107-140
4. Lin, T.Y.: Generating concept hierarchies/networks:mining additional semantics in relational data. Advances in Knowledge Discovery and Data Mining, Proceedings of the 5th Pacifi-Asia Conference, Lecture Notes on Artificial Intelligence 2035 (2001) 174-185
5. Lin, T.Y.: Granular computing, LNAI 2639 (2003) 16-24
6. Ma, J.-M., Zhang, W.-X., Li,T.-J.: A covering model of granular computing, Proceedings of 2005 International Conference in Machine Learning and Cybernetics (2005) 1625-1630
7. Pawlak, Z.: Rough sets, International Journal of Computer and Information Science 11 (1982) 341–356
8. Pawlak, Z.: Rough sets, Kluwer Academic Publishers. Dordrecht Boston London (1991)
9. Skowron, A., Stepaniuk, J.: Information granules: towards foundations of granular computing, International Journal of Intelligent Systems 16 (2001) 57–85
10. Yao, Y.Y.: Granular computing: basic issues and possible solutions, Proceedings the 5th joint conference on information sciences (2000) 186-189
11. Yao, Y.Y.: Information granulation and rough set approximation, International Journal of Intelligent systems, 16 (2001) 87-104
12. Yao, Y.Y.: A partition model of granular computing, LNCS 3100 (2004) 232-253
13. Yao, J.T., Yao, Y.Y.: Introduction of classification rules by granular computing, LNAI 2475 (2002) 331-338
14. Zadeh, L.A.: Fuzzy sets and information granularity. In: Advances in Fuzzy Set Theory and Applications, Gupta, N, Ragade, R and Yager. eds. NorthHolland, Amsterdam (1979) 3-18
15. Zadeh, L.A.: Fuzzy logic-computing with words. IEEE Transactions on Fuzzy Systems, 4 (1996) 103-111
16. Zadeh, L.A.: Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems **19** (1997) 111-127
17. Zhang, W.-X., Wu, W.-Z., Liang, J.-Y., Li, D.-Y.: Theory and Methods of Rough Sets, Science Press, Beijing (2001)