

General set approximation and its logical applications*

Zoltán Ernő Csajbók^a, Tamás Mihálydeák^b

^aDepartment of Health Informatics,
Faculty of Health, University of Debrecen,
csajbok.zoltan@foh.unideb.hu

^bDepartment of Computer Science,
Faculty of Informatics, University of Debrecen
mihalydeak.tamas@inf.unideb.hu

Abstract

To approximate sets a number of theories have appeared for the last decades. Starting up from some general theoretical pre-conditions the authors give a set of minimum requirements for the lower and upper approximations and define general partial approximation spaces. Then, these spaces are applied in logical investigations. The main question is what happens in the semantics of the first-order logic when the approximations of sets as semantic values of predicate parameters are used instead of sets as their total interpretations. On the basis of defined partial interpretations, logical laws relying on the defined general set-theoretical framework of set approximation are investigated.

Keywords: Approximation of sets, rough set, partial logic, partial semantics

MSC: 03B52, 03B60, 68T37

1. Introduction

Classical set-theoretic approximation space was invented by Zdzisław Pawlak in the early 1980's. It is known as rough set theory [9, 10]. From the beginning, its many generalizations have been developed [1, 12, 18]. In this paper, first, minimum requirements for set approximations relying on general theoretical pre-conditions are given and a general approximation framework is defined. This scheme is allowed

*The publication was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

to treat uniformly the common features of classical rough set theory and its different generalizations. Next, these results are applied in logical investigations.

2. General approximation spaces

The starting point of rough set theory is a nonempty finite set U and an equivalence relation ε on U . The equivalence classes are called ε -*elementary sets*. Any unions of ε -elementary sets are called *definable*. A subset $S \in 2^U$ can be naturally approximated by the lower and upper ε -approximations of S which are denoted by $\underline{\varepsilon}(S)$ and $\overline{\varepsilon}(S)$, respectively. The former is the union of all ε -elementary sets which are the subsets of S , whereas the latter is the union of all ε -elementary sets which have a nonempty intersection with S .

In rough set theory, the equivalence classes form a partition of U , i.e., they are pairwise disjoint and cover the universe. Therefore, the Pawlakian theory of set approximations can be generalized by giving up the pairwise disjoint property and/or the covering of the universe.

Giving up the disjoint property but retaining the covering, a natural generalization of rough set theory is obtained which is called the covering-based rough set theory [1, 18]. The partial nature of real-life problems, however, requires working out partial approximation schemes. Thus, not only the pairwise disjoint property but also the covering of the universe are given up. This basically new approach is referred to as general (partial) approximation of sets [5, 6, 4].

Let U be a finite nonempty set. The set-theoretical framework of the general approximation of sets has the following components:

- *domain*, not necessarily a finite set derived from U whose members are approximated;
- *base sets*, some beforehand detached sets in the domain forming a *base system*;
- *definable sets* deriving from base sets as possible approximations of sets in the domain (base sets are always definable);
- *approximation pair* determining lower and upper approximations of sets.

There are many approximation space types mainly depending on how deriving the domain from ground set and definable sets from base sets, in addition how approximation pair determining the lower and upper approximations.

At the beginning of the investigations, we stipulate the following two fundamental pre-conditions:

- lower and upper approximations of any set must be definable;
- lower approximation of any set must be included in its upper approximations.

The intuitive meaning behind this scheme is the following. The background knowledge about objects of interest is represented by the base system. Base sets

and definable sets can be considered as primary and derived tools in the approximation process, respectively. Derived tools or simply tools constitute all available knowledge about the objects. They are used to describe approximately any sets belonging to the domain. After all, lower and upper approximations must be definable sets so that we are able to make any decision relying on tools.

Definition 2.1. The 5-tuple $\text{GAS}(U) = \langle U, \mathfrak{B}, \mathfrak{D}_{\mathfrak{B}}, l, u \rangle$ is a *general approximation space* with the domain 2^U if

1. $\mathfrak{B}(\neq \emptyset) \subseteq 2^U$ and if $B \in \mathfrak{B}$, $B \neq \emptyset$ (*base system*);
2. $\mathfrak{B} \subseteq \mathfrak{D}_{\mathfrak{B}} \subseteq 2^U$ and $\emptyset \in \mathfrak{D}_{\mathfrak{B}}$ (*definable sets*);
3. lower and upper approximation mappings $l, u : 2^U \rightarrow 2^U$ form an ordered pair $\langle l, u \rangle$, called a *weak approximation pair*, with the following minimum requirements:
 - $l(2^U), u(2^U) \subseteq \mathfrak{D}_{\mathfrak{B}}$ (l, u are *definable*);
 - if $S_1 \subseteq S_2$ ($S_1, S_2 \in 2^U$), $l(S_1) \subseteq l(S_2)$, $u(S_1) \subseteq u(S_2)$ (*monotonicity*);
 - $u(\emptyset) = \emptyset$ (*normality* of u);
 - $l(S) \subseteq u(S)$ ($S \in 2^U$) (*weak approximation property*).

Definition 2.2. $\text{GAS}(U)$ is *total*, if the base system \mathfrak{B} covers the universe, i.e., $\bigcup \mathfrak{B} = U$ and *partial* otherwise.

If $B \in \mathfrak{B}$ is a union of a family of sets $\mathfrak{B}' \subseteq \mathfrak{B} \setminus \{B\}$, B is called *reducible* in \mathfrak{B} , otherwise B is *irreducible* in \mathfrak{B} . A base system \mathfrak{B} is *single-layered* if every base set is irreducible, and *one-layered* if the base sets are pairwise disjoint.

Definition 2.3. $\text{GAS}(U)$ *relies on Pawlakian base* if \mathfrak{B} is a partition of U .

Corollary 2.4. $\text{GAS}(U)$ *relies on Pawlakian base if and only if its base system is total and one-layered.*

The weak approximation property immediately implies the following statement.

Corollary 2.5. *In the general approximation space $\text{GAS}(U)$, $l(\emptyset) = \emptyset$ also holds (normality of l).*

It is reasonable that base sets as primary tools are exactly approximated from “lower side”. In certain cases, it also holds for definable sets.

Definition 2.6. A weak approximation pair $\langle l, u \rangle$ is

1. *granular* if $l(B) = B$ ($B \in \mathfrak{B}$);
2. *standard* if $l(D) = D$ ($D \in \mathfrak{D}_{\mathfrak{B}}$).

It is an important question how lower and upper approximations relate to the approximated set itself.

Definition 2.7. A weak approximation pair $\langle l, u \rangle$ is

1. *lower semi-strong* if $l(S) \subseteq S$ ($S \in 2^U$), i.e., l is contractive;
2. *upper semi-strong* if $S \subseteq u(S)$ ($S \in 2^U$), i.e., u is extensive;
3. *strong* if it is lower and upper semi-strong, i.e., each subset $S \in 2^U$ is bounded by $l(S)$ and $u(S)$: $l(S) \subseteq S \subseteq u(S)$.

Corollary 2.8. In the general approximation space $\text{GAS}(U)$,

1. if $\langle l, u \rangle$ is an upper semi-strong approximation pair then $u(U) = U$ (co-normality of u);
2. if $\langle l, u \rangle$ is an upper semi-strong approximation pair and l is standard, then $l(U) = U$ (co-normality of l).

Definition 2.9. The general approximation space $\text{GAS}(U)$ is a *weak/standard/lower semi-strong/upper semi-strong/strong approximation space*, if the approximation pair $\langle l, u \rangle$ is weak/ standard/lower semi-strong/upper semi-strong/strong, respectively.

Definition 2.10. In the general approximation space $\text{GAS}(U)$, let us define $\mathfrak{D}_{\mathfrak{B}}$ with the following inductive definition:

- $\emptyset \in \mathfrak{D}_{\mathfrak{B}}$, $\mathfrak{B} \subseteq \mathfrak{D}_{\mathfrak{B}}$;
- if $D_1, D_2 \in \mathfrak{D}_{\mathfrak{B}}$, then $D_1 \cup D_2 \in \mathfrak{D}_{\mathfrak{B}}$.

The approximation space of this type is called *strictly union type*.

Definition 2.11. Let $\text{GAS}(U)$ be a strictly union type approximation space, and let us define the lower and upper approximations as follows:

1. $l(S) = \bigcup \mathbb{L}(S)$, where $\mathbb{L}(S) = \{B \in \mathfrak{B} \mid B \subseteq S\}$ ($S \in 2^U$);
2. $u(S) = \bigcup \mathbb{U}(S)$, where $\mathbb{U}(S) = \{B \in \mathfrak{B} \mid B \cap S \neq \emptyset\}$ ($S \in 2^U$).

Then, $\langle l, u \rangle$ is called the *Pawlakian approximation pair*.

A strictly set-union type approximation space with a Pawlakian approximation pair is called a *Pawlakian approximation space*.

Corollary 2.12. If $\text{GAS}(U)$ is a Pawlakian approximation space, it is lower semi-strong and granular.

3. Logical system based on set approximation

Rough set theory can be considered as the foundation of various kinds of deductive reasoning. In particular, various kinds of logics based on the rough set approach have been investigated, rough set methodology contributed essentially to modal logics, many valued logic, intuitionistic logic (see in [11]). There are many papers about the logical features of different systems of rough sets (or, in general, set approximation). A summary of this research can be found in [14].

In the last years we presented some papers dealing with the possibilities of using different systems of set approximation in logical semantics. At RST2011 a very general framework of set approximation was presented and we showed that the semantics of a partial first-order logic could rely on general partial approximation spaces.¹ The common set theoretical framework proved a useful tool to compare the results and consequences of different approximations from the logical point of view.

The theoretical results appear on the meta level of our logical system. Our starting point is a given language of first-order logic and a finite distinguished subset of its predicate parameters. Its members express available concepts/properties and relations which are called tools.

According to a usual interpretation of the given first-order language, the interpretation of all predicate parameters is obtained. Then, the following question appears. What does it have to be changed if we use the approximations of sets generated by the semantic value of our tools as semantic values of predicate parameters instead of the sets given by their total interpretation?

Our solution regarding this question is that we may introduce three different partial interpretations of the given first-order language with the help of its total interpretation. As a result, the “real world” that appears in the total interpretation and its approximations that appear in generated partial interpretations can be compared. If we take into consideration all interpretations that fulfill some requirements, we have the logical possibility to investigate what happens to logical laws when the approximations of sets instead of the sets themselves are used. The common set theoretical framework proves to be a useful tool to compare the results and consequences of different approximations from the logical point of view.

Having given the language of a logical system, a general partial approximation space is generated by the help of an interpretation of the language in order to give semantic rules. Finally central semantic notions are defined in order to give some fundamental laws.

Later we showed that

1. there is a logically exact way to define approximative functors on object level in order to determine what kind of approximation has to be taken into consideration in the evaluating process of a formula;
2. the representations of concepts (properties) of our available knowledge can be used to approximate not only any concept (property) but any relation.

In order to represent approximative functors in object language a specific first-order language is required. Its main reason is that in standard first-order language there is no predicate functors, i.e., functors whose inputs and outputs are predicates. There are two different types of approximative functors: the first ones produce predicates from predicates, and the second ones produce formulae from

¹http://rst.disco.unimib.it/RoughSetTheory/Slides2011_files/8-Csajbok-Mihalydeak.pdf

formulae. The first ones can be treated as primitives, the second ones can be introduced by contextual definitions. Informally these functors tell us what kind of approximations (lower or upper) has to be used in order to determinate the truth value of a given formula.

Having a tool-based interpretation of a language of tool-based partial first-order logic (TbPFoL), the semantic values of tools (the members of set \mathcal{T}) determine a general partial approximation space with respect to the given interpretation. The generated approximation space is logically relevant in the sense, that it gives the lower and upper approximations of any predicate P to be taken into consideration in the definition of semantic rules.

For example if P is a one-argument predicate parameter which is not a tool and $u \in U$, then

1. P^∇ is true/false at u if our tools evaluate P as certainly true/false at u ;
2. P^∇ is undefined at u if our tools are not enough to decide whether P is certainly true or certainly false at u ;
3. P^Δ is true/false at u if our tools evaluate P as maybe true/certainly false at u ;
4. P^Δ is undefined at u if our tools are not enough to decide whether P is maybe true or certainly false at u .

From the logical point of view, the main advantage of our logical framework is the flexibility. It can be recognized on different levels:

1. The generated partial interpretations rely on two theoretical points:
 - (a) The set of semantic values of tools is given by the total interpretation $\langle U, \varrho \rangle$. These semantic values represent available knowledge, i.e., the total interpretation gives us the representations of available concepts/properties and relations by which we approximate any concept/property or relation (with respect to the given interpretation).
 - (b) The general approximation space is generated by tools with respect to the given interpretation. By specifying an approximation pair $\langle l, u \rangle$, different theoretical systems of approximation can be gained: Pawlak-type, generalized Pawlak-types, approximations which are used in different systems of granular computing or in very general versions of the approximation of sets.
2. In the definition of one consequence relation, the different notions of models can be used. For instance, we can require that all lower models of the set of premises should be a lower/upper/mixed model of the conclusion: it makes possible to investigate different approximations comparably.

References

- [1] BONIKOWSKI, Z., BRYNIARSKI, E., WYBRANIEC-SKARDOWSKA, U., Extensions and intentions in the rough set theory, *Information Sciences* **107**(1-4), 149–167 (1998)
- [2] DÜNTSCH, I., GEDIGA, G., Approximation operators in qualitative data analysis, In: de Swart, H.C.M., Orłowska, E., Schmidt, G., Roubens, M., eds.: *Theory and Applications of Relational Structures as Knowledge Instruments*. Vol. 2929, LNCS, Springer, 214–230 (2003)
- [3] CSAJBÓK, Z., Partial approximative set theory: A generalization of the rough set theory, In: Martin, T., Muda, A.K., Abraham, A., Prade, H., Laurent, A., Laurent, D., Sans, V., eds.: *Proceedings of the International Conference of Soft Computing and Pattern Recognition (SoCPaR 2010)*, December 7-10, 2010., Cergy Pontoise / Paris, France, IEEE, 51–56 (2010)
- [4] CSAJBÓK, Z.E., Approximation of sets based on partial covering. In: Peters, J.F., Skowron, A., Ramanna, S., Suraj, Z., Wang, X. (eds.) *Transactions on Rough Sets XVI*. LNCS, vol. 7736, pp. 144–220. Springer-Verlag, Berlin, Heidelberg (2013)
- [5] CSAJBÓK, Z., MIHÁLYDEÁK, T., A general set theoretic approximation framework. In: Greco, S., Bouchon-Meunier, B., Coletti, G., Fedrizzi, M., Matarazzo, B., Yager, R.R. (eds.) *Proceedings of IPMU 2012, Catania, Italy, July 9-13, 2012, Part I*. CCIS, vol. 297, pp. 604–612. Springer-Verlag, Berlin Heidelberg (2012)
- [6] CSAJBÓK, Z., MIHÁLYDEÁK, T., Partial approximative set theory: A generalization of the rough set theory. *International Journal of Computer Information Systems and Industrial Management Applications* **4**, 437–444 (2012)
- [7] LIN, T.Y., Approximation theories: Granular computing vs rough sets. In Chan, C.C., Grzymala-Busse, J.W., Ziarko, W., eds.: *Rough Sets and Current Trends in Computing*, 6th International Conference, RSCTC 2008, Akron, OH, USA, October 23-25, 2008, Proceedings. Vol. 5306, LNCS, Springer, 520–529 (2008)
- [8] MIHÁLYDEÁK, T., Partial first-order logic with approximative functors based on properties. In: Li, T., Nguyen, H.S., Wang, G., Grzymala-Busse, J., Janicki, R., Hassanién, A.E., Yu, H. (eds.) *Rough Sets and Knowledge Technology*. 7th International Conference, RSKT 2012, Chengdu, China, August 17-20, 2012, Proceedings. LNAI, vol. 7414, pp. 514–523. Springer-Verlag, Berlin Heidelberg (2012)
- [9] PAWLAK, Z., Rough sets. *International Journal of Computer and Information Sciences* **11**(5), 341–356 (1982)
- [10] PAWLAK, Z., *Rough Sets: Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Dordrecht (1991)
- [11] PAWLAK, Z., SKOWRON, A., Rudiments of rough sets. *Information Sciences* **177**(1) 3–27 (2007)
- [12] PAWLAK, Z., SKOWRON, A., Rough sets: Some extensions. *Information Sciences* **177**, 28–40 (2007)
- [13] PAWLAK, Z., POLKOWSKI, L., SKOWRON, A., Rough sets: An approach to vagueness. In Rivero, L.C., Doorn, J., Ferraggine, V., eds.: *Encyclopedia of Database Technologies and Applications*. Idea Group Inc., Hershey, PA, 575–580 (2005)
- [14] POLKOWSKI, L., *Rough Sets: Mathematical Foundations*. Advances in Soft Computing, Physica-Verlag, Heidelberg (2002)

-
- [15] SKOWRON, A., STEPANIUK, J., Information granules: Towards foundations of granular computing. *International Journal of Intelligent Systems* **16**(1), 57–85 (2001)
 - [16] SKOWRON, A., ŚWINIARSKI, R., SYNAK, P., Approximation spaces and information granulation. *LNCS Transactions on Rough Sets* **3400**(3), 175–189 (2005)
 - [17] YAO, Y.Y., Information granulation and rough set approximation. *International Journal of Intelligent Systems* **16**(1), 87–104 (2001)
 - [18] YAO, Y.Y., On generalizing rough set theory. In: Wang, G., Liu, Q., Yao, Y., Skowron, A. (eds.) *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, 9th International Conference, RSFDGrC 2003, Chongqing, China, May 26-29, 2003, Proceedings. *LNAI*, Vol. 2639, pp. 44–51. Springer-Verlag, Berlin Heidelberg (2003)
 - [19] ZHU, P., Covering rough sets based on neighborhoods: An approach without using neighborhoods. *International Journal of Approximate Reasoning* **52**(3), 461–472 (2011)
 - [20] ZADEH, L.A., Granular computing and rough set theory. In Kryszkiewicz, M., Peters, J.F., Rybinski, H., Skowron, A. (eds.) *Rough Sets and Intelligent Systems Paradigms*, International Conference, RSEISP 2007, Warsaw, Poland, June 28-30, 2007, Proceedings. Vol. 4585, *LNCS*, Springer, 1–4 (2007)