Volume 34 Number 4 (114)

Janusz A. POMYKAŁA

Wyższa Szkoła Menedżerska w Warszawie, Wydział Informatyki Stosowanej i Systemów Bezpieczeństwa

VALUATION AND APPROXIMATION OF INFORMATION TECHNOLOGY PROJECTS

Summary. We consider information technology projects in the methodology of PMI described with the approach of Kathy Schwalbe. We recall formal approaches to knowledge representation. We analyze approximation spaces of Pawlak and Żakowski introducing several operations and we suggest applications to the design of projects. We introduce several methods to evaluate project, also idea to evaluate expert estimations concerning projects is expressed.

Keywords: lattice, fixed point, information systems, technology projects

OCENIANIE PROJEKTÓW TECHNOLOGII INFORMACYJNEJ

Streszczenie. W artykule stosujemy metodologię PMI budowy projektów w ujęciu Kathy Schwalbe. Przypominamy formalne podejścia do problemu reprezentacji wiedzy. Opisujemy własności operacji przybliżających i sugerujemy zastosowanie do oceny projektów systemów informacyjnych.

Słowa kluczowe: przybliżanie, kraty, punkt stały, system informacyjny, technologia informacyjna

1. Introduction

The notion of project in information technology is very important. Projects are realized often using technology of relational databases or object oriented databases [5,14]. However equally important is the problem how to design complex and expensive projects. This problem is related to modern system analysis, project management, to behavioral sciences, to the theory of organization and methods of optimization.

To develop theory connecting notions of roughness with information technology, we will need knowledge from several domains: Topology [3], Logic [2, 15, 25], Boolean Algebras[3], Databases [14], Data mining [24]. In definitions and concepts related to projects we follow the book of Kathy Schwalbe [1].

Main idea of the paper is the following: how can we apply ideas of approximation, especially in term of rough sets, to the design of and to valuation of projects (especially information technology projects). We will describe evaluation of projects on several levels of abstraction: Approximation and evaluation of projects; and Evaluation of expert decisions, valuations. In other words, on every level of designing and working on the project, and working on the information technology system, we can approximate the final shape of the system by several steps (or by discerning several levels of approximation).

Next idea is the following: having estimation or valuation of several projects, we can compare not only projects but also the levels of knowledge expressed by the persons.

Approximation operations defined by Professor Z. Pawlak in approximation space based on indiscernibility relations, are well examined. Generalizations of these operations in relational systems and in covering spaces are examined in the frame of covering rough sets. In last year (2010, 2011, 2012) many papers have been devoted to this subject. I will mention here about some of them.

2. Project

A **project** is set of activities to accomplish a unique purpose, organization objectives, tasks, improvements. [1].

It is very important to understand what is the main **purpose** of the project, what are main aims, tasks, activities, jobs. In information technology projects key product might be new software or patent; even if the group of people will work for a period of months on the subject, if they will not get the final result, project can not be closed.

Projects have the following attributes [1]:

- Unique purpose a well-defined objective.
- Requires resources from various areas including people, hardware, software and other assets.
- Time project has definite beginning and definite end; how long should it take to complete the project.
- Scope what unique product or service will be expected from the project.
- Sponsor project should have a primary sponsor or customer.

• Involves uncertainty – it is difficult to clearly define the project's objectives, cost and time:

It is possible that uncertainty is inherently connected to data, functions or relations; in that case we will use theory of fuzzy sets or the theory of rough sets to describe properly the situation or to solve a problem. Let me add also the following attributes:

• Logic:

Logical organization-which tasks are more important or have dependencies on the other tasks? Here we can also mean functional dependencies in relational database and its generalizations-multivalued and inclusion dependencies (see eg. [2], [5], [28]).

- Importance –projects involving new technologies are challenging, the management is difficult, cost may be high and risk quite big.
- Value many projects give work for the society, produce goods, create new ways of living, types of thinking, solve problems, build new environments, create new ideas, which can change the society.
- Extent let us imagine worldwide help for poor people, or continental scientific aerospace projects; let us consider projects concerning water on the Earth or soil on the continent.

Now to estimate or valuate the project, we have to know at least percentage values for some of the attributes above and also qualitative values for other attributes.

It is important to meet scope and time goals, support organizational objectives, prepare well work breakdown structure and focus on satisfying project stakeholders and sponsors [1]. Equally important is to have methods for dealing with uncertainty, risk, uncertain knowledge, cyber-hacking, control and improvement or cyber-terrorism.

In Project Management Institute the following elements of the methodology are designed (see [1]):

- 1. Project management process groups.
- 2. Integration management.
- 3. Scope management.
- 4. Time management.
- 5. Cost management.
- 6. Quality management.
- 7. Human resources.
- 8. Communication management.
- 9. Risk management.
- 10. Procurement management.

Approximation operations are used for points 3, 4, 5, 8 and 9.

In this article we describe information technology projects in terms of approximation space, using several notions of this kind of space. We shall begin from definitions.

Valuation of project

We shall give several definitions concerning valuation of project. First we formulate "local" definition based on the view of project as the sequence of tasks, which must be done to close the project. It is simple idea – calculation of the project valuation in this case is just weighted sum. The more challenging is the definition in which we consider more complex projects, with possibly hierarchical structure and when we need estimation of tasks expressed in different data structures. In this case we have uncertain knowledge or attributes with uncertain values.

Having valuation of projects at our disposal, we can compare valuations given by experts, as a consequence we are able to check which one of the experts is more precise, clever, deeply thinking?

Definition 1

Let T1,...,Tn are the tasks which have to be done in the project P. Let Per(Tk) denotes percent of the finished work to close task k. Let wk denotes weight associated with the task Tk (or with the class of the similar tasks). The following real number will be called valuation of the project:

$$Z=(Per(T1)*w1+...+Per(Tn)*wn)/n$$
(1)

Now let us assume that expert 1 estimates project by value x, and expert 2 estimates project value by y.

If |Z-x| < |Z-y| then expert 1 better estimates value of the project P then expert 2.

In the more general case let us assume that we have m technology projects with weights c1...cm describing the complexity of each project. We also assume that expert 1 estimates projects giving values v1...vm, and expert 2 estimates them giving values w1...wm. Then weighted sum e1= c1*v1+...+cm*vm represents global valuation of projects by expert1 and e2= c1*w1+...+cm*vm represents global estimation given by expert2. If we assume that there exists other objective function estimating values of projects, then the distance from e1, e2 to this function, expresses how close to the proper valuations are values given by experts.

Defnition 2

Let SC1,...,SCn denotes scope of the project divided on n-areas (or regions). If SCI is realized we give the value 1 for this area, if SCI even not started, we give value 0; otherwise we give value between 0 and 1 proportional to the level of finishing the scope area. Arithmetical mean of these values is defined to be scope valuation of the project.

Example

Let us assume that experts valuate several attributes of the project (Table 1).

Attributes of the project

Table 1

	Expert1	Expert2	Expert3	AVG
Appropriate to ols	1/3	1/2	2/3	
Appropriate skills	1/2	3/5	2/5	
Efficient methods	1	1/2	1	
Clear solutions	1	1	3/4	
Accurate	2/3	3/4	4/7	
Well organized	1	1	1	
Complete	0.6	0.7	0.8	

If all values belong to the set [0,1], then by definition valuation of the part of the project described by the above attributes is just average of all values. We can also find minimal and maximal values of the project given by experts 1,2,3.

Now to estimate experts we define function:

$$\sum |(\text{Exi}(\text{Attr})-\text{AVG})| \tag{2}$$

Here sum is over all attributes. Expert is valuated to be one of the best from a set of experts evaluating the project, if the above function has minimal value.

Example

Table 2

Example			
	Ex1	Ex2	
Project mgt.	Well	Good	
Integration mgt.	very well	Fine	
Scope mgt.	Almost	Not good	
Time mgt.	Quite well	Quite good	
Cost mgt.	Best	Very good	

Total value given by Ex1 is 1/5*(3+5+2+4+6) and by Ex2 is 1/5*(4+4,5+3+3,5+5).

3. Approximation

In 1980 polish mathematician professor Z. **Pawlak** introduced new kind of information system and related notions of approximation space and rough sets. The space was based on the equivalence relations and on interpretation of indiscernibility relation. Later some authors (see [2-30]) and many others, generalize the above concepts by introducing similarity relations, tolerance relations and binary relations, approximation based on coverings, concrete and abstract approaches and axiomatic systems. On the abstract level, propositions presented

below may be expressed or at least interpreted in view of lattice theoretical fixpoint theorems of B. Knaster and A. Tarski.

Definition 1

Let us assume that universal set of elements of interest is given, this set will be denoted by U. U is also called the universe of objects. The family of sets C will be a covering of U indexed by the set T, if $C=\{C_t: t \text{ in } T\}$ and the union of the family C is equal to U.

Definition 2

For X being arbitrary subset of U we define:

LX is the union of all elements C_t of the covering C which are included in the set X; This set is called lower approximation of the set X; Operation L is called lower approximation operation.

HX is the union of all elements C_t from the covering C which have nonempty intersection with the set X. H is called upper approximation operation on the set U defined with respect to the covering C.

Example 1

Let us assume that we consider algebra of intervals on the real line. The elements of the algebra are half closed intervals [a,b), such that a<b, b is real number or infinity. Let us form the union of all intervals [a,b) such that a>0. It is lower approximation of the set $(0,\infty)$. On the other hand the union in the algebra is half interval $[0,\infty)$.

Definition 3

The following four tuple (U, C, L, H) will be called **Żakowski** space, where U is the universe of objects, C is a covering of U, i.e. the union of the family C is equal to U, L is lower approximation operation, H is an upper approximation operation.

Remark

In general operations L and H are not conjugated which means that -L(-X) does not have to be equal to H(X), for X a subset of U. Equivalently -H(-X) can be not equal to L(X).

Fixed points of approximation operations are called definable sets.

Definition 4

The covering of U is a partition if the elements of it are disjoint. Usually partition is denoted by P.

If C=P is a partition of U, the space (U, P, L, H) is called **Pawlak's approximation space**, where L, H are approximation operations defined with respect to the partition P. Let me also define **indiscernibility neighborhood** of element x with respect to the covering C:

$$O(x,C)=U\{K \in C : x \in K\}$$
(3)

For any element x in U the set

 $I(x,C) = I(x) = \{y \in U: \text{ for every } Ct \text{ in } C \text{ holds}(x \text{ is in } Ct \text{ iff } y \text{ is in } Ct)\}$ is called **iniscernibility** kernel of x with respect to the covering C.

Remark: the notation in my original papers was slightly different, $O_x^{\ C}$ has been used for indiscernibility neighborhood of x, and ${\bf I_x}^{\ C}$ denoted **kernel** of x, with respect of the covering ${\bf C}$.

Basic properties:

Lemma 1

Every equivalence relation in U determines partition on universe U. Every partition on U determines equivalence relation on universe U. Every Pawlak's space is Żakowski's space.

Lemma 2

Suppose (U, C) is an approximation space. The following conditions are equivalent

$$H(H(X))=H(X)$$
 for every $X \subseteq U$

- 1. $H(H(\lbrace x \rbrace))=H(\lbrace x \rbrace)$ for every x in U
- 2. $\{O_x^C : x \text{ in } U\}$ is a partition of the set U.

Corollary

1. The operation H is a closure operation iff the family $\{ O_x^{\ C} : x \text{ in } U \}$ is a partition of the set U.

Now we define the lower operation which is dual (conjugated) to the operation H:

Definition

The operation L_1 given by

$$L_1(X) = \{x: \forall K \in C(x \in K \Longrightarrow K \subset X\}$$

will be called a weak lower approximation operation.

Theorem:([9])

The following conditions are equivalent:

- 1. $\{O_x^C : x \text{ in } U\}$ is a partition of the set U.
- 2. H is a closure operation.
- 3. $L_1(X) = \{x: \forall K \in C(x \in K \Rightarrow K \subseteq X) \text{ is an interior operation. Moreover topologies generated by operations H, L₁ are identical.$

Now we examine the lower operation L. Let C be a covering of the set U.

Lemma

C is a subbase of a topology on U.

C is a base of the topology on U if C is closed on finite intersections of the sets belonging to C.

Theorem ([9])

The operation L_1 is the interior operation iff the covering C is the base of the topology generated by C.

Corollary

If C is not the base of the topology generated by subbase C, then L_1 is not interior operation.

If the operation L_1 is not interior operation then C is not the base of the topology generated by subbase C.

In other words the operation L_1 of the lower approximation is characteristic for coverings C being proper (i.e. being not partitions). This shows in some sense the difference between Pawlak and Żakowski spaces.

Definition 5

If X is a subset of U and LX is not equal to HX then X is called a **rough** set. Minimal subcovering of C is called a **reduct** of the covering C.

Definition 6

Chain of sets with respect to intersection will be linearly ordered finite family of sets with the property that every set and its successor have nonempty intersection. In a similar way infinite chain of sets is defined; ϵ -chain contains sets with the diameter smaller than epsilon. **Component** of X in C is by definition family of all finite chains with the property that one of elements of the chain intersects X.

For A, B subsets of the universe U, Com(A,B) relation holds iff A,B are in the same component in covering C

Definition 7

Assume x,y belong to U and C is a covering of U. x, y are **inseparable** iff there is no E belonging to C s.t. one of elements x,y belongs to E and the other doesn't. We define relation **Ker** with respect to covering C in U as follows:

x Ker y iff x,y are inseparable in space (U,C).

Ker is called kernel relation.

Definition 8

The space (U, C, Ker, Com) will be called *general covering space*. Let L(Ker) denotes lower approximation operation with respect to partition Ker. Let H(Ker) denotes upper approximation w.r.t. Ker. L(Com), H(Com) are defined analogically w.r.t. partition Com.

The tuple (U, C, Ker, Com, L(C), H(C), L(Com), H(Com), L(Ker), H(Ker)) will also be called general covering space.

Remark

Definition is adequate because relations Com, Ker are equivalence relations; (by abuse of notation Com, Ker are also partitions of U).

By <= shall be denoted inclusion relation.

Lemma2

Every equivalence relation in U defines the partition of U Every partition in U defines the equivalence relation on U $L(Ker)X \le L(C)X \le L(Com)X$.

$$H(Ker)X \le H(C)X \le H(Com)X$$

4. Axiomatization

One of the important problems for approximation operations is abstract axiomatization. First paper in which abstract characterization for Pawlak operations was found, is 1994 paper of T.Y. Lin and Qing Lin:

Theorem (Lin, Liu):

For the pair of rough operators $X \rightarrow (H(X), L(X))$ which satisfy the axioms

- 1. $H(A \cup B) = H(A) \cup H(B)$
- 2. *L*(*A*)*⊆A*
- 3. $H(\emptyset) = \emptyset$
- 4. L(L(A))=L(A)
- 5. L(U-X)=U-H(X)
- 6. L(X)=H(L(X))

Then there is equivalence relation R such that for every X

$$H(X)=H(R)(X)$$
 And $L(X)=L(R)(X)$.

For covering lower approximation operation the axioms are given by W. Zhu, F.Y. Wang:

Theorem (Zhu, Wang):

If an operation $L:P(U) \rightarrow P(U)$ satisfies the following properties: for any $X,Y \subseteq U$

- 1. L(U)=U
- 2. $X\subseteq Y$ implies $L(X)\subseteq L(Y)$
- 3. $L(X) \subseteq X$
- 4. L(L(X))=L(X)

Then there exists a covering C of U, such that the covering lower approximation operation L_C generated by C equals to L.

5. Application

Now, if we have values of attributes not belonging to reals or natural numbers, we assign to every attribute special scale on which we can estimate values and then give them order. Valuation of one such scale can be further counted in specific way, for example finding AVG or GVG, or

- a) compared with scale for other attribute, and being the base for the decision which attribute is better, best or smaller, bigger etc.
- b) If attribute expresses knowledge of the expert, then we can infer who of the expert is more clever, who is the proper designer for the given information system, etc.
- c) If values of attributes belong to sets included in fuzzy sets, rough sets, membranes or in other specific sets of values, we should use respective algebra. For example for rough sets we can use Stone algebra[12], Boolean algebra[4], or relations algebra, covers and other structures [20-30].

6. Lattice theoretical fixed point theorem

In this section we use an elementary fixed point theorem which holds in arbitrary complete lattices. In 1927 Bronisław Knaster and Alfred Tarski proved set theoretical fixpoint theorem by which every function, on and to the family of all subsets of a set, which is increasing under set theoretical inclusion, has at least one fixpoint. In 1939 Alfred Tarski proved a lattice theoretical fixed point theorem. We shall formulate this theorem and we show some consequences in approximation space.

Definition

By a **lattice** we understand a system $[A, \leq]$ formed by a nonempty set A and a partial order \leq in A, such that for every two elements there exist greatest lower bound and also least upper bound.

The lattice is called **complete** if every subset B of A has a least upper bound UB and a greatest lower bound $\cap B$.

The function f on B to C, where B,C are subsets of A, is increasing if, for any elements x,y, $x \le y$ implies $f(x) \le f(y)$. By a **fixpoint** of a function f we understand an element of the domain of f such that f(x)=x.

Theorem (Tarski)

Let

- i. $\langle A, \leq \rangle$ be a complete lattice
- ii. f be an increasing function on A to A,

P be the set of all fixed points of the function f.

Then the set P is not empty and the system $\langle P, \leq \rangle$ is a complete lattice.

In particular we have

 $UP = U\{x: f(x) \ge x\} \in P$

and

 $\bigcap P = \bigcap \{x: f(x) \leq x\} \in P$.

Corollary 1

If H is upper approximation operation in Pawlak space then every union of indiscernibility classes is fixed point of H.

If L is lower approximation operation in Pawlak space then every union of indiscernibility classes is fixed point of L.

If B is union of indiscernibility equivalence classes then it is fixed point of H and L.

If H is upper approximation operation in Pawlak space then every union of indiscernibility classes is fixed point of H

Corollary 2

If **L** is lower approximation operation in Żakowski space, then its fixed points are exactly sets which are unions of elements of the covering C.

Corollary 3

If **H** is upper approximation operation in Żakowski space, then its fixed points are exactly sets which are unions of components of the covering C.

Definition

Fixed points of H and L in Pawlak space are called **definable** sets.

If F is operation in a space S then every fixed point of F is called definable with respect to operation F in the space S.

Acknowledgement

I am grateful for Professors-Adam Obtułowicz, Marek Greniewski and Jacek Pomykała for valuable discussion about information systems.

BIBLIOGRAPHY

- 1. Schwalbe K.: Managing Information Technology Projects. Course Technology, 2010.
- 2. Demri S. P., Orlowska E. S.: Incomplete information: Structure, Inference, Complexity. Springer, 2002.
- 3. Engelking J.: Topologia. PWN, Warszawa 1970.
- 4. Iwinski T. B.: Algebraic approach to rough sets. Bull. Pol. Ac. Math. 35, 1987.
- 5. Maier D.: The Theory of Relational Databases. Comp. Sc. Press, 1983.
- 6. Obtulowicz A.: Mathematical models of uncertainty systems. Natural Computing 2, 251 263, Kluwer 2003.
- 7. Paun Gh.: Computing with Membranes. Journal of Computer and System Science 61, 2000, p. 108÷143.
- 8. Pawlak Z.: Information Systems, Theoretical Foundations. Inf. Systems 3, 1981, p. 205÷218.
- 9. Pomykala J. A.: Approximation Operations in Approximation Space. Bull. Pol. Ac. Math., Vol. 35, No. 9÷10, 1987.
- 10. Pomykala J. A.: Some Remarks on Approximation. Demonstratio Mathematica, Vol. XXIV.
- 11. Pomykala J. A., Pomykala J. M.: On regularity of hypergraph sequences. Demonstratio Mathematica, Vol. XXVII, No. 3÷4.
- 12. Pomykala J. A., Pomykala J. M.: The Stone algebra of rough sets. Bull. Pol. Ac. Math., 1988.
- 13. Skowron A., Polkowski L.: Rough sets in knowledge discovery. Springer, Physica-Verlag, Heidelberg 1998.
- 14. Ullman J. D.: Principles of database systems. Computer Science Press, 1980.
- 15. Vakarelov D.: Modal logics for knowledge representation systems. TCS 90, Elsevier, 1991, p. 433÷456.
- 16. Wille R.: Restructuring lattice theory: an approach based on hierarchies of concepts, [in:] Rival I. (ed): Ordered sets, Reidel. Dordrecht-Boston 1982, p. 445÷470.
- 17. Żakowski W.: Approximations in the space {U, Π). Demonstratio Math. 16, 1983.

- 18. Samanta P., Chakraborty M. K.: Consistency-Degree Between Knowledges, [in:] Kryszkiewicz M. et.al. (eds.): RSEISP 2007.
- 19. Pomykała J. A.: On similarity based approximation of information. Demonstratio Mathematica, Vol. XXVII, No 3÷4, 1994.
- 20. Rasiowa H., Skowron A.: Approximation Logic. Mathematical methods of specification and synthesis of software systems, Academie Verlag, Band 31, Springer, Berlin 1986.
- 21. Pomykała J. A.: About tolerance and similarity relations in information systems. Rough sets and Current trends in computing, LNCS LNAI 2475, Springer, 2002.
- 22. de Haas E., Pomykała J. A.: A Note on Categories of Information Systems, [in:] Ziarko W. P. (ed.): Rough sets, fuzzy sets and knowledge discovery. Springer, 1994, Workshops in computing: Banff, Alberta, Canada.
- 23. Wiweger A.: On topological Rough Sets. Bulletin of the Polish Academy of Sciences, Mathematics, Vol. 37, No. 1÷6, 1989, p. 43÷54.
- 24. Cios K. J., Pedrycz W., Swiniarski R. W.: Data mining methods for knowledge discovery. Kluwer Academic Publishers, 1998.
- 25. Salwicki A., Mirkowska G.: Logika Algorytmiczna dla programistów, WNT, Warszawa 1992.
- 26. Zhu W., Wang F. Y.: Reduction and axiomatization of covering generalized rough sets. Inf. Sc. 152, 2003, p. 217÷230.
- 27. Greco S., Hata Y., Hirano S., Inuiguchi M., Miyamoto S., Nguen H. S., Słowiński R. (eds.): Rough Sets and Current Trends in Computing. LNAI 4259, Springer, 2006.
- 28. Alpigini J. J., Peters J. F., Skowron A., Zhong N. (eds.): Rough Sets and Current Trends in Computing. LNAI 2475, Springer, 2002.
- 29. Kryszkiewicz M., Peters J. F., Rybinski H., Skowron A. (eds.): Rough Sets and intelligent Systems Paradigms. LNAI 4585, Springer, 2007.
- 30. Ziarko W. P. (ed.): Rough sets, fuzzy sets and knowledge Discovery. Springer, 1994, Workshops in computing: Banff, Alberta, Canada.

Omówienie

W artykule omawiamy metodę oceny projektów technologii informacyjnej na podstawie metod zarządzania projektami określonymi w książce Kathy Schwalbe [1], zgodnie z metodologią PMI. Każdy etap budowy systemu może być oceniony osobno, a następnie określa się ocenę łączną, zwykle zdefiniowaną jako sumę ważoną. Niektóre z etapów mogą być przybliżane za pomocą operacji aproksymacyjnych. Operacje te zostały ustalone przez wielu autorów, m.in.: Z. Pawlaka, W. Żakowskiego, U. Wybraniec-Skardowską, J. A. Pomykałę, A. Nakamura, A. Skowrona, M. Chakraborty, M. Banjerjee, A. Obtułowicza, Y. Y. Yao,

W. Zhu, T. J. Li, T. B. Iwińskiego, Z. Bonikowskiego, T. Bryniarskiego i wielu innych. Na poziomie bardziej abstrakcyjnym formalizm aproksymacji może być wyrażony za pomocą twierdzenia o punkcie stałym Tarskiego-Knastera. Poruszamy tylko wstępnie problem aksjomatycznej charakteryzacji wybranych operacji przybliżających.

Address

Janusz Andrzej POMYKAŁA: Wyższa Szkoła Menedżerska w Warszawie, ul. Kawęczyńska 36, Warszawa, Polska, janusz.andrzej.pomykala@outlook.com, janusz.pomykala@mac.edu.pl.